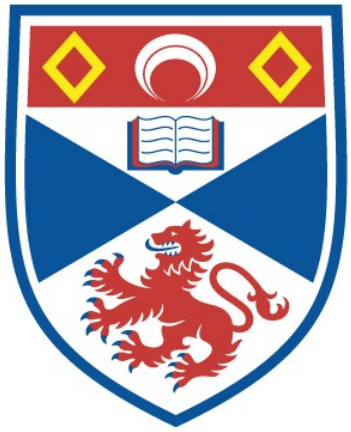
HUMPBACK WHALES IN BRAZIL: DISTRIBUTION, ABUNDANCE AND HUMAN IMPACTS

Guilherme Augusto Bortolotto de Oliveira

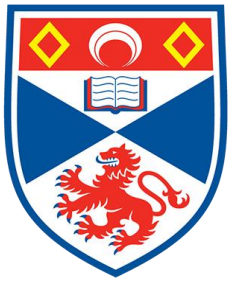
A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



2019

Humpback whales in Brazil: distribution, abundance and human impacts

Guilherme Augusto Bortolotto de Oliveira



University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of

Doctor of Philosophy (PhD)

at the

University of St Andrews

January 2019

Intentionally left blank.

Candidate's declaration

I, Guilherme Augusto Bortolotto de Oliveira, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 50,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree.

I was admitted as a research student at the University of St Andrews in March 2015.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date Signature of candidate

Supervisors' declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date Signature of supervisor

Date Signature of supervisor

Date Signature of supervisor

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Guilherme Augusto Bortolotto de Oliveira, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis):

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

Date Signature of candidate

Date Signature of supervisor

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Guilherme Augusto Bortolotto de Oliveira, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date Signature of candidate

ACKNOWLEDGEMENTS

To my wife, Manoela, and to our son, Ícaro: thank you for the happy and *full-of-love* life that we live together.

To my parents, sister, brother and family, and to my uncles: thank you for supporting and inspiring me.

To my supervisors, Phil Hammond, Len Thomas and Alex Zerbini: thank you for all your patience, support and lessons. When I look back to the beginning, I realise how much more I learned from you than how to become a better researcher. Especially to Alex, who believed in me when I needed an opportunity and, most importantly, kept believing. This has made a huge difference.

I am very thankful to some friends, who are researchers, and who contributed to essential steps of my formation: Álvaro Menin, for the first opportunity in research; Daniel Danilewicz, for supervising my masters and for all the amazing opportunities; Artur Andriolo, for the continuous inspiration; Cristiane Kolesnikovas, for all the support, amazing professional example and for the ongoing partnership; Ubirajara Costa, for the trust and support.

Nadya and Joyce: without you, things would have been much harder during my PhD. Thank you for everything.

To my office mates, and attached-office mates, in the Bute: Saana, Alex, Anna, Ben, Claire, Crystel, Eilidh, Emily, Filipa, Iosu, Jamie, Janneke, Kagari, Luis, Mathilde, Miguel, Nadya, Patrick, Paul, Sara, Takashi, Tilen and Tomoko: thank you for making this a great place to work.

To all the brilliant people from CREEM and SMRU, especially Andy, Aubrie, Camille, Carl, Catriona, Charles, Chris M., Chris McK., Clint, Danielle, Dave B., Dave M., Dave T., Debbie, Eric, Gabriel, Heather, Izzy, Jeremy, Jo, Joe, Kelly, Lars, Laura, Lindsay, Louise, Lucia, Marie, Mark, Matt C., Matt B., Michail, Monica, Monique, Phil, Popi, Rene, Rhona, Rick, Ryan, Sascha, Simon, Sonja, Sophie, Steve Balfour, Steve Buckland,

Theo, Theoni, Tiago, Valentin: thank you for all the amazing lessons and for all the cake. Both things made me grow a lot.

I am thankful for the long time friendship of Lucas Farias, André Rizzo, Fran, Igor, Carol, Mari Neves, Lucas M., Fernanda, Caio, Eliza and Leo.

The Monitoring Whales by Satellite Project (*Projeto Monitoramento de Baleias por Satélite*, PMBS) research cruises were sponsored by Shell Brasil. The Universidade Federal de Rio Grande (FURG) and the R/V Atlântico Sul crew provided essential support during fieldwork.

I am grateful to several staff members at the University of St Andrews who provided me with excellent conditions to study.

Debbie Russell and Vincent Ridoux: thank you for reviewing this thesis and for providing valuable feedback on my work.

This work was supported by the Brazilian National Council for Scientific and Technological Development (*Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq, Programa Ciência sem Fronteiras*, scholarship 208203/2014-1).

TABLE OF CONTENTS

ABSTRACT

The humpback whale population breeding in Brazilian waters has greatly increased, after near extinction due to whaling in the twentieth century. Today, these animals are under pressure from human activities in the area. In this thesis, habitat use, distribution, abundance, population status and potential threat from oil spills were investigated to improve knowledge of the population's ecology and provide useful information for conservation and management. Distribution and habitat use were investigated through spatial models applied to line transect data and to tracking data. Line transect data were also used to estimate abundance. Distribution maps from both data types were used with a simulation of oil dispersion to evaluate risk of impact from oil spills. The new abundance estimates, together with information on population increase from another study, were used to update a Bayesian population dynamics model to re-assess population status. Modelling of line transect data indicated whale density to be higher in slower currents, at shorter distances to both the coastline and shelf edge, and at sea surface temperatures between 24 and 25°C, and to be related to shelter. A higher concentration of animals was predicted in the southern portion of Abrolhos bank, an enlargement of the continental shelf. Modelling of tracking data agreed with those findings, despite differences in the nature of the data and analytical methods. Risk maps of oil spill impact indicated that areas in the south of the breeding range present highest risks to the animals. Abundance estimates (14,264, CV = 0.084, in 2008; 20,389, CV = 0.071, in 2012) provide further evidence that the population is increasing, and contributed to improved precision in the population status assessment. New information provided here will inform conservation of the humpback whale population breeding in Brazilian waters and the need for, and implementation of, any necessary management action.

Intentionally left blank

CHAPTER 1

GENERAL INTRODUCTION

The distribution of animals and their use of space are driven by biological priorities (predator avoidance, prey availability, competition, etc.) according to specific needs of individuals, populations or species, all which may vary with time (Matthiopoulos and Aarts, 2010). Moreover, the heterogeneity in use of space by animals is influenced by requirements for specific resources, which are often not homogeneously distributed in the environment (Matthiopoulos, 2003). Distribution may be described as "the product of factors that act in a parallel or interactive way over different scales of space and time on each species, and sometimes on groups of species" (Forcada, 2009), and can also be expected to vary according to specific requirements of the individuals, that may be related to age, sex, health status, size, etc. In summary, factors influencing animal distribution may be physical, biological, anthropogenic, demographic, ecological, evolutionary, or a combination of these.

In marine mammal species, especially cetaceans, distribution has mainly been studied in relation to oceanographic characteristics and processes including: depth, sea bottom slope, distance to physical features (shore line, isobaths, banks, islands, etc.), water temperature, salinity, productivity and sea currents, with many examples of such studies. Hamazaki (2002) predicted cetacean species occurrence in the mid-western Atlantic Ocean from sea surface temperature (SST), monthly front probability (which was derived from a different source of SST), and ocean-bottom depth and slope. Pirotta et al. (2011) found that topographic features were strongly related to sperm whale (*Physeter macrocephalus*) distribution, with water temperature also closely linked to differences in patterns of space-use among solitary individuals and groups. For Hector's dolphins (*Cephalorhynchus hectori*) around Banks Peninsula, east coast of New Zealand, Rayment et al. (2010) found that their occurrence was strongly related to distance from the shore. Depth had a seasonal effect and this feature might be linked either to the seasonal distribution of their prey or to the shift of particular features that pregnant females probably prefer during the summer period, such as sheltered waters. Pardo et al. (2015)

showed that the absolute dynamic topography (ADT) was a good predictor of the distribution of blue whales (*Balaenoptera musculus*) in the north-eastern Pacific Ocean, ADT being strongly linked to the distribution of their main prey, krill. The above examples illustrate that animal distribution can be related to fixed and dynamic features of the environment, and that when studying this topic, model descriptions of complex animal-environment relationships are unavoidably a simplification of reality.

The differential usage of space is often treated as a proxy for habitat preference (Matthiopoulos, 2003), so the environmental variables that are measured should have logical links to biological features. In most circumstances, the biological feature most likely to influence marine mammal distribution is their prey. Although the distribution, abundance and, especially, availability of prey species are difficult to measure, they are likely to be influenced by the properties of the ocean, such as the concentration of chlorophyll (Hastie et al., 2005; Dalla Rosa et al., 2012) and other productivity indices determined by physical characteristics of the water column (Bradshaw et al., 2004).

Predator avoidance may also influence habitat use of some marine mammals (Wirsing et al., 2008). To improve survival chances, it is expected that animals as prey should avoid areas of higher risk of predation (Lima, 1998). Heithaus and Dill (2006) showed that bottlenose dolphins (*Tursiops aduncus*) in Shark Bay, Western Australia, shift their foraging sites according to fluctuation in tiger shark (*Galeocerdo cuvier*) densities in the area. The dolphins foraged in areas of reduced productivity, with fewer fish for them to prey on, but with better chances of escaping in the event of encountering sharks.

Human presence and activities are also likely to influence marine mammal distribution (Sorensen et al., 1984), since they are part of the environment experienced by animals. Acoustic deterrent devices (ADDs) deployed around fish farms have shown to cause pinnipeds (Götz and Janik, 2013) and cetaceans (Nowacek et al., 2007) to avoid part of their usual habitat, at least temporarily. Other studies have documented the impact of pile-driving for installation of offshore wind farms on porpoises in the North Sea, with animals temporarily avoiding areas previously occupied (e.g., Gilles et al., 2009; Scheidat et al., 2011; Dähne et al., 2013).

Motivated by the potential future development of hydrocarbon extraction/production activities in the South-eastern Bering Sea, tracking data were used to study the movements of North Pacific right whales (*Eubalaena japonica*) relative to a protected area, and to regions of interest for human activities (Zerbini et al., 2015). Conclusions in that work support that a "Critical Habitat" (in U.S. terminology) area likely comprises the most important habitat for this population, but that its summer range overlaps fishery activities and shipping routes.

For marine mammals that migrate seasonally, such as most baleen whale species, there is a clear difference between the factors expected to drive the use of space at each seasonal site (Corkeron and Connor, 1999). Specifically for humpback whales (*Megaptera novaeangliae*), social organization seems to have a major influence on the distribution of groups in their breeding areas (Martins et al., 2001; Ersts and Rosenbaum, 2003). Presence of calves within groups and reproductive status where found to be related to the distribution of this species in Hawaiian waters (Craig and Herman, 2000). Humpback whale competitive groups (Clapham et al., 2009), which are surface active groups with many animals generally displaying aggressive behaviour related to reproduction, are more frequently found in relatively deeper waters than other groups (Clapham, 2000). In contrast, in feeding areas their distribution is well documented to be influenced by habitat features related to prey distribution, such as bathymetric features and productivity (e.g., Dalla Rosa et al., 2012).

1.1 IMPACTS OF HUMAN ACTIVITIES ON MARINE MAMMALS

Humans may negatively affect marine ecosystems in several different ways, from removal of animals, either deliberately or indirectly, to climate change and cumulative impacts of multiple threats (Crain et al., 2009; Pirotta et al., 2018). In a global scale study, Halpern et al. (2008) concluded that the world's oceans are entirely affected by humans, with about a third of marine ecosystems suffering very strong impacts. Only regions towards the poles remained relatively unaffected, which can be expected to change in the near future since those areas are the most affected by climate change (McMillan et al., 2015). The threats posed by anthropogenic activities on many marine mammal populations around the world are the main motivations for interest in their conservation.

The killing of animals is clearly the most evident and direct way that human activities can impact marine mammals. Historically, millions of whales were removed by whaling over several centuries (Rocha et al., 2014) and whaling during the twentieth century has been considered the greatest exploitation of wildlife in human history (Clapham et al., 2009). Whaling is very much reduced today and it is not a present threat to the majority of the populations, compared to other pressures. Instead, many cetacean species are impacted by fishery activities, and bycatch is recognized as being responsible for killing a large number of animals in different areas around the world (e.g., Hofman and Bonner, 1985; Secchi et al., 2002; Read et al., 2006). For example, bycatch is the main threat to the franciscana (*Pontoporia blainvillei*), due to its highly restricted area of occurrence and marked preference for estuarine and very shallow waters, areas which are under severe influence from human disturbance (Danilewicz et al., 2010). The vaquita porpoise (*Phocoena sinus*), with its very restricted range, is facing imminent extinction. Vaquitas suffer with bycatch (Taylor et al., 2017) in nets targeting the also rare totoaba (*Totoaba macdonaldi*), a fish sought after for its swimming bladder, an expensive delicacy in parts of Asia. Large whales can also be affected by entanglement and concerns exist for some coastal populations, such as right whales in the North Atlantic (*Eubalaena glacialis;* Knowlton and Kraus, 2001) and humpback whales in the Gulf of Maine (Robbins and Mattila, 2004). Lethal consequences of large whale entanglement are not rare (e.g., Knowlton and Kraus, 2001; Cassoff et al., 2011; Meÿer et al., 2011).

Increasing vessel traffic is a global conservation concern for marine environments, with large and small cetaceans being among the most vulnerable animal species (Van Waerebeek et al., 2007). The physical presence and noise produced may affect the distribution of marine mammals (Erbe, 2002; Cartwright et al., 2012) with possible consequences for reproductive success (Sousa-Lima and Clark, 2009). Although the majority of ship-strikes on marine mammals probably still remain unreported, they are likely to be responsible for hundreds of deaths every year (Laist et al., 2001). Studies to assess the risk of collisions between cetaceans and ships have shown the importance of managing marine traffic in particular areas or time periods (e.g., Kite-Powell et al., 2007; Williams and O'Hara, 2010; McKenna et al, 2012; Bezamat et al., 2015).

Pollution in the marine environment may impact species in a number of ways, including by ingestion of debris (Bortolotto et al., 2016d) and entanglement (Laist, 1997), and potentially contribution to the development of various kinds of diseases, due to immunosuppression (Alonso et al., 2012; Kajiwara et al., 2004; Dorneles et al., 2015). Marine mammals are at a relatively high trophic level, being particularly susceptible to diseases, because immune suppression may be caused by bioaccumulation of chemical pollutants (Handoh and Kawai, 2014). Environmental pollution can also affect how cetaceans distribute in space. Higher densities of bottlenose dolphins (*Tursiops truncatus*) in the Gulf of Olbia, Sardegna, Italy, were found in areas relatively more impacted by human activities within their distribution range (López, 2018). Findings in that study suggest that the dolphins were attracted by either increased food availability as result of aquaculture activities (e.g., entangled fish and fish farming waste) or by indirect increase of prey attracted to the area due to coastal pollution.

Another conservation concern for some species is the effect of anthropogenic noise, which can be generated by military activities, seismic surveys, construction in the marine environment, coastal development, tourism and generally the noise from boat propellers/engines. These sources of noise can interfere with communication, disturb normal behaviour or cause trauma in the hearing system (Richardson and Würsig, 1997). Nowacek et al. (2007) reviewed the documented reactions of marine mammals in the presence of anthropogenic noise, concluding that one of the most commonly observed behavioural responses is displacement.

Besides the direct impact that human activities in the marine environment may have on individuals, cumulatively they may also have broader consequences and effects for the entire population (King et al., 2015; Pirotta et al., 2018). For example, northern elephant seal (*Mirounga leonina*) pup survival and population growth rates may have been reduced due to alteration in foraging behaviour in response to the presence of human activities in Macquarie Island, Australia (New et al., 2014). The level of impact depends on the biology of the species and Bailey et al. (2014) discussed that the constrained foraging area of harbour seals (*Phoca vitulina*) in Scotland during their breeding season makes the animals vulnerable to human influence, which could negatively impact reproduction. Croll et al. (2001) also found that noise from commercial ships may influence the

distribution, residency and foraging behaviour of large whales, during sound exposure experiments in San Nicolas Island, California, U.S.

Oil spills are a potential threat to many kinds of marine life, as demonstrated by consequences of oil accidents in the past (Fraser, 2014). Oil spill models (Kileso et al., 2014) can combine characteristics of the environment (wind, current, temperature, etc.) and of the oil (density, volume etc.) to predict the destination of oil drifting in the sea. The results of such models can potentially help to identify areas of higher risk for marine species in case of a future accident, if these are analysed together with information on species distribution.

1.2 MARINE MAMMAL DISTRIBUTION STUDIES TO INFORM CONSERVATION AND MANAGEMENT

Marine mammals usually influence several trophic levels (Pauly et al., 1998) and provide important ecological services for the environment (Roman et al., 2014). Therefore, their conservation is necessary for naturally balanced ecosystem maintenance. Conservation actions for marine mammal populations, and management strategies of human activities, need information from ecological studies. For conservation of species in environments where conditions are constantly changing, it is important to have the best understanding of the interactions between key species and their environment (Víkingsson et al., 2015). Identifying areas important to these animals and how they use them are among the most desirable information.

As an example, the feeding grounds of the humpback whale population that breeds along the coast of Brazil were described in Zerbini et al. (2006), through the use of satellite telemetry tracking of whales. Their feeding grounds are now confidently considered to be in offshore waters off South Georgia and the South Sandwich Islands, and the previous hypothesis that this population uses the Antarctic Peninsula during its feeding season (Slijper, 1962) has been rejected. Because of that information, today it is known that whaling in South Georgia and the South Sandwich Islands severely affected the Brazilian population, and that high latitude waters also need conservation in order to permit this population to recover.

A common way to protect cetacean populations is through implementing Marine Protected Areas (MPAs; Hoyt, 2012). Species distribution models can be strong tools to identify core areas and support such implementation (Cañadas et al., 2005; Redfern et al., 2006), and can also be used to inform mitigation of human activities that may affect species. For example, Bailey and Thompson (2009) developed species-habitat models to predict areas of higher density of bottlenose dolphins (*Tursiops truncatus*), harbour porpoises (*Phocoena phocoena*), grey seals (*Halichoerus grypus*) and harbour seals. The authors found that the Special Area of Conservation (SAC) designated under the EU Habitats Directive to promote the conservation of bottlenose dolphins was also effective at protecting the other species because of the overlap in distribution in the area, and because human activities such as fishing are restricted within the SAC. In British Columbia, Canada, spatial models have been used to predict critical habitat for large whales, considering a few environmental variables (Gregr and Trites, 2001). Results support that the area is important for sperm whale reproduction annually during April and May, and that sheltered bays and straits along the coast form the preferred habitat for humpback whales. Embling et al. (2010) predicted harbour porpoise distribution off the west coast of Scotland and then used a criterion that is a trade-off between maximum habitat protection and minimum costs of managing the protected area (Possingham et al., 2000) to suggest areas most suitable for implementing of SACs. Bailey et al. (2014) integrated different sources of tracking data on harbour seals in Scotland to characterize an offshore region as an important and consistently used foraging area. Several other studies of marine mammal distribution, habitat use, abundance and population trends have suggested management actions in order to conserve species, such as the expansion of existing marine protected areas and changes in human activities to avoid impact (e.g., Wilson et al., 1997; Baumgartner and Mate, 2005; Cañadas et al., 2005; Ward et al., 2011; Dalla Rosa et al., 2012; Castro et al., 2014). An important consideration is that, mainly for migratory species with demarked breeding and feeding seasons (e.g., humpback and right whales), models developed to understand and predict distribution specifically for feeding areas are unlikely to provide accurate predictions for breeding areas, and vice-versa, because the animals' priorities and needs are different at each place (Corkeron and Connor, 1999; Redfern et al., 2006). As another example of how biological information can be used to influence management, seismic and other oil and gas exploitation activities

on the Abrolhos Bank, which is an important humpback whale breeding habitat in Brazil, were cancelled during their breeding season (winter-spring) after a study on the potential impact of oil operations in the area suggested such activities could impact the animals (Engel et al., 2004; Marchioro et al., 2005).

1.3 METHODS TO STUDY MARINE MAMMAL DISTRIBUTION

Matthiopoulos and Aarts (2010) summarized the main sources of data and methods used for studying the spatial distribution of marine mammals, which are primarily from line transect surveys and telemetry tracking of individuals. Individual recognition data (e.g., photo-identification), which have characteristics of both survey and tracking data, and opportunistic sightings may be used, and even whaling data have been used to derive distribution models (Gregr and Trites, 2001; Torres et al., 2013). Also, when associated with drifting prediction models, strandings data may be useful for estimating cetacean distribution. For example, Peltier et al. (2014) used a drift model to infer the distribution of common dolphins from dead animals found stranded along the coast of the British Channel and Bay of Biscay. If standardized stranding monitoring procedure are followed, this type of data can be very informative and can parallel other methods, such as sighting surveys (Peltier et al., 2014). Analytical methods to study marine mammal distribution usually model the use of space as a function of environmental variables, which typically represent proxies for habitat preference or habitat use, as mentioned above (Matthiopoulos, 2003). Environmental variables may be obtained during the survey (e.g., measures of water temperature from a boat, while searching for animals), derived from available databases (e.g., depth from nautical charts), generated from models of oceanic processes (e.g., currents dynamic, modelled SST) or derived from remotely sensed data (Redfern et al., 2006). Often, the sources of environmental covariates to a study are those that provide the best temporal and spatial resolution, precision and expected accuracy desired in each case.

A typical way of obtaining data to study the distribution of cetaceans is through line transect surveys (Buckland et al., 2001; Miller et al., 2013), from surface and/or aerial platforms. Animals or groups of animals are detected while searching along transect lines and the data are used to generate estimates of density. Not all animals/groups will be detected, so an important issue is the correct estimation of the detection probability, which

may be influenced by characteristics of the survey platform (e.g., platform height and speed), the sighting conditions (e.g., sea conditions as measured on the Beaufort scale) and features directly related to the animal or group of animals (e.g., sighting cue or group size) (Hammond, 2010). Depending on the scale considered, spatial and temporal correlation of detection locations need to be dealt with in the models (Matthiopoulos and Aarts, 2010), so the model residuals are not correlated. Independence of model residuals is a target in every modelling exercise (Wood, 2017) and several considerations must be made, such as regarding sample size. In general, the more sightings data available, the better habitat use and distribution can be inferred. In that context, Virgili et al. (2018) pooled data from several line transect surveys, to model the distribution of deep-diving cetaceans in the north Atlantic (beaked whales, sperm whales and kogiids). Because of their cryptic behaviour and usually long dives, sightings for those species are relatively rare, and the pooling allowed that study to considerably improve the modelling of habitat use.

Tracking data collected via animal-borne devices, which may use different systems for obtaining locational data (e.g., GPS, Argos satellite system, GLS) are typically used in studies of pinnipeds but also in studies of large cetaceans (Panigada et al., 1999; Bailey et al., 2014; Cerchio et al., 2016). Such devices can be archival (i.e., devices need to be recovered for assessing data) or transmit the data through telemetry, when data are sent remotely (i.e., via Argos system, GSM, UHF). Tracking data may provide information about vertical as well as horizontal position and behaviour, which are not easily obtained from other techniques (Bailey et al., 2014). As mentioned above, unlike transect sampling, these data collection methods provide information from sampling individual animals rather than space. Several common issues exist in modelling space/habitat use from tracking data, including dealing with the variation in space accessibility, the bias from tagging locations, unbalanced data, variable correlation, and non-constant detection rate (Aarts et al., 2008). The major problem is spatiotemporal autocorrelation in the data because consecutive observations come from sequential locations in space and in time (Aarts et al., 2008; Matthiopoulos and Aarts, 2010). However, this correlated structure is useful when deriving corrected locations from locations observed with error and/or inferring animal behaviours from tracking data (Jonsen et al., 2005).

Data collection and modelling procedures are continuously evolving; studies presenting different data sources to study distribution of the same population, such as telemetry tracking and line transect methods, may thus provide a good opportunity for a more complete understanding of how animals are dispersed in their habitats. Moreover, the comparison of information obtained from sampling either space (i.e., in line transect surveys) or individuals (i.e., from tracking animals), may help to better understand distribution and habitat use, and to evaluate the efficacy of different methods on several aspects.

A common aim of modelling the use of space by animals is to identify high use areas by predicting animal distribution to the entire area of interest from the data collected in the portion that was sampled (Redfern et al., 2006). To develop prediction models, data for the predictor variables must be available for the whole study area and not only for the places sampled during survey effort. This may limit the range of variables to be used; some of those that are found to not be available may be those expected to influence animal distribution, and the use of alternative features linked to that may be necessary (Hamazaki, 2002). Prediction of habitat of highly mobile ocean species may not require very detailed information on predictor variables, depending on the scale that is chosen, the scope and extent of the study (Redfern et al., 2006). However, in breeding areas and generally in small scale studies, animals may respond to fine scale habitat variability (e.g., shallow and sheltered areas; Ersts and Rosenbaum, 2003; Torres et al., 2013). In summary, inferences and conclusions must always match temporal and spatial scales considered in the studies. Another aim of modelling habitat use may be to identify areas potentially suitable to be occupied for a species of population (Redfern et al., 2006); predictions outside the study area are strongly reliant on the assumption that the model is a good description of the patterns observed in the data, and most importantly, the modelled covariate-response relationships are assumed to be applicable to that area. Such approach can be useful to investigate cetacean distribution over a large scale and in poorly documented areas. For example, Mannocci et al. (2015) used aerial line transect survey data to model and extrapolate densities of cetaceans to far beyond the regions surveyed in that study. Having surveyed several tropical areas, cetacean density was extrapolated to the entire circumference of the planet, but care was taken to not predict distribution to areas where environmental features departed from the ranges within the surveys.

Methods typically used to analyse the relationships between marine mammals and their habitat are based on regression models. Relationships between marine mammal occurrence or density and habitat variables are not naturally expected to be linear and most studies have used Generalized Additive Models (GAMs; Hastie and Tibshirani, 1990; Wood, 2017), which fit smoothing functions to accommodate non-parametric relationships between response and predictor variables. Mainly because of their flexibility and robustness, GAMs are probably the most useful method currently applied to model marine mammal distribution (Matthiopoulos and Aarts, 2010). When residual autocorrelation is present, which is generally difficult to avoid for tracking data, Generalized Estimation Equations (GEEs) or mixed models may be useful tools (Fieberg et al., 2009; Bailey et al., 2013). Development of analytical techniques to deal with other issues that data present when modelling marine mammal distribution include improvement of smoothing to better fit the variability in the data (Scott-Hayward et al., 2017), modelling locations in space as "point processes" (Renner et al., 2015), and considering animal tracks in a grid over the study area as a result of a Markov process (Whitehead and Jonsen, 2013). As one example of uncommonly applied method, Bailey et al. (2014) used a Bayesian state-space approach to integrate tracking data from harbour seals in Scotland from multiple sources. Their study, besides characterizing an important foraging area, provided some improvement of the analytical techniques.

1.4 HUMPBACK WHALES ALONG THE BRAZILIAN COAST

Every year, from winter to late spring, the Brazilian coast is inhabited by humpback whales that use the area for reproduction (Martins et al., 2001; Andriolo et al., 2010). Calving and mating occur in these low latitude tropical waters, while feeding occurs around South Georgia and the South Sandwich Islands in the summer (Zerbini et al., 2006; Engel and Martin, 2009). Whales concentrate over the Abrolhos Bank, an enlargement of the continental shelf where about 80% of the population is expected to be found every breeding season (Andriolo et al., 2010). During that time, their typical range is over the continental shelf (i.e., from shore to shelf break; Figure 1.1), from Natal (~5^{o}S), north-eastern, to Cabo Frio (~23^{o}S), south-eastern Brazil (Zerbini et al., 2006). These animals are considered by the International Whaling Commission to be part of the Breeding Stock A, BSA, (IWC, 2005).

The Brazilian coast was an important whaling ground in the past, with whaling stations distributed from the north-eastern to the southern shore-lines. Pre-modern catches date from 1603, when the main target was the southern right whale (Edmundson and Hart, 2014; Morais et al., 2017). With the collapse of the right whale population in the area around the beginning of the twentieth century, humpback whales were targeted because they also have a consistent coastal distribution (Morais et al., 2017). This time was also marked by the transition from artisanal to modern whaling methods, including factory and steam-powered ships and explosive harpoons. Modern whaling methods resulted in annual catches increasing significantly, with the BSA population being exploited to near extinction in the 1950s (Zerbini et al., 2011). Since their almost total protection in the late 1960s, the population has been recovering (Andriolo et al., 2010; Ward et al., 2011; Zerbini et al., 2011; Pavanato et al., 2017; Wedekin et al., 2017). The expected consequences of the increase in the population on interactions between animals and human activities in the area remain unclear. Castro et al. (2014) found that MPAs along the Brazilian coast do not effectively provide protection to humpback whales in their wintering habitat. Major concerns arise from the increasing interest in oil and gas extraction in the region, which is set to expand in the near future (BRASIL, 2015).

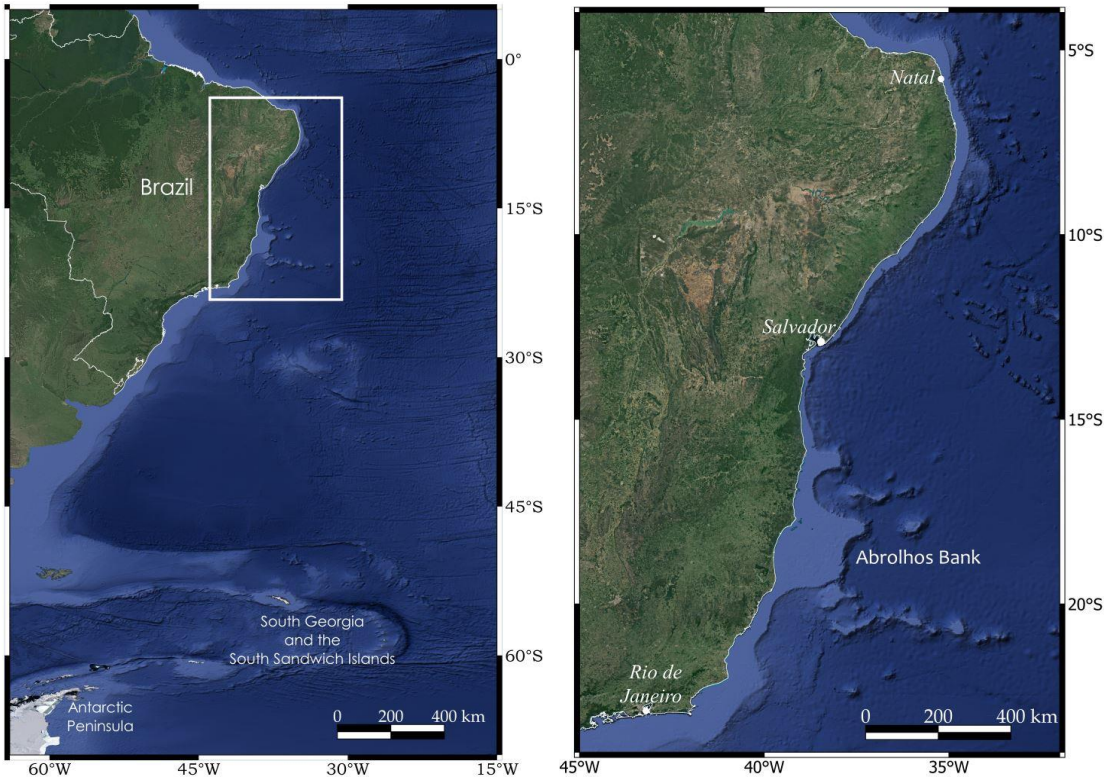


Figure 1.1. Distribution range of humpback whales in the western south Atlantic. The map on the right shows the typical range of whales during the breeding season.

Studies of humpback whales along the coast of Brazil include abundance estimation using mark-recapture (Freitas et al., 2004), line transect sampling from aerial (Andriolo et al., 2006; 2010; Pavanato, 2017) and ship surveys (Zerbini et al., 2004; Bortolotto, 2016a; b), descriptive habitat-use (Martins et al., 2001; Rossi-Santos et al., 2008) and site fidelity (Wedekin et al., 2010). A recent study focused on their population growth rate (Wedekin et al., 2017) and indicates that the population is not only increasing, but very rapidly for this species and those with similar life-histories. Satellite telemetry tracking of animals has also been conducted (Zerbini et al., 2006), which described migration routes and sites of departure of animals leaving the area at the end of the breeding season.

The distribution of humpback whales at their breeding sites should be influenced mostly by reproductive priorities (Craig and Herman, 2000; Martins et al., 2001; Ersts and Rosenbaum, 2003; Cartwright et al., 2012). However, variables that relate to this may be difficult to obtain at a large scale for humpback whales. For example, Ersts and Rosenbaum (2003) discuss the presence of calves in humpback whale groups as

influencing the differential distribution of groups in part of a breeding area. For that study, a closing mode approach (Hammond, 2010), in which the group is approached after being detected from a designed survey line, was adopted, permitting more detailed observations. This procedure would increase substantially the time needed to cover the area of interest for humpback whales sin Brazil.

Notwithstanding the different hypotheses proposed to explain large whale migration (Corkeron and Connor, 1999), spending the colder winter months in the area seems likely to convey physiological and ecological benefits for the animals, especially for calves, by reducing heat loss. The absence of humpback whales in their feeding areas during winter also coincides with the seasonal increase of the ice layer in Antarctic waters that is a key factor for southern krill reproduction (Nicol, 2006). This permits krill to take advantage of the high productivity of the region, in the absence of some predators, to reproduce and form the huge and dense shoals that are them exploited by the whales, when they return in the summer.

Besides a better understanding of the factors influencing humpback whale ecology in a highly impacted environment, the availability of different sources of data (i.e., telemetry tracking and line transect) to study their distribution represents a unique opportunity to compare different analytical methods and results generated from sampling animals (telemetry tracking) and space (line transect); this can also be used to better evaluate potential risks of impact from human activities. In addition, a better understanding of the population's present conservation status can inform the need to implement any management actions, which should consider current knowledge of habitat use, distribution and potential risk of impacts for these animals in the area.

Investigating population status is essential to inform the need for conservation. For the population of humpback whales off Brazil, Zerbini et al. (2011) applied a density-dependent, sex and age-aggregated population dynamics model (Pella and Tomlinson, 1969) to data on catch, growth rate and abundance, to estimate several parameters related to population dynamics. That study adopted a backward approach (Butterworth and Punt, 1995), starting from a recent abundance estimate in 2005 to back project the population size at the beginning of 20th century, which is considered in the model as the carrying capacity of the population. Similar procedures, or more complex models, have been used

by the Scientific Committee of the International Whaling Commission to improve understanding of the status of several whale species (e.g., Punt and Butterworth, 1999; Wade, 2002; Jackson et al., 2006; Brandon et al., 2007; Johnston et al., 2011).

Updating population assessment models with recent data can provide guidance on the need for implementing management strategies and/or conservation actions currently or in the future, and how urgent they might be. That, associated with robust information on population distribution and potential impacts, allow such important decisions to be better informed.

1.5 THESIS OBJECTIVES

This PhD thesis has the objective to improve understanding of the ecology of humpback whales inhabiting the Brazilian coast, their distribution, habitat use, threats from oil spills to the population and its conservation status.

Chapter 2 investigates the distribution, habitat use and abundance of humpback whales along the Brazilian coast from line-transect data, as a function of environmental variables. Sightings data were available from two research cruises that surveyed the area occupied by humpback whales in Brazil: 416 sightings in 2008; 553 in 2012. Spatial models were fitted to explain how whale abundance and distribution were related to environmental variables; new model-based estimates of abundance were computed and compared with the existing design-based estimates.

Chapter 3 investigates the distribution and habitat use of humpback whales in their area of occurrence in Brazilian waters from tracking data, as a function of environmental variables, to compare to the outputs from Chapter 2. The data available were from more than 100 whales that were tagged with satellite transmitters from 2003 to 2012 in the area. Spatial models were fitted to tracking data to investigate distribution and habitat use, and the results were compared with those from the previous chapter.

Chapter 4 evaluates the risk of oil spills impacting humpback whales in Brazilian waters. A simple and fast method is presented to identify areas were oil spilled could overlap with the distribution of whales. The method uses a simulation of oil trajectory at the sea surface

and information on whale distribution predicted in Chapter 2 and 3 to identify areas of potential higher risk of impact to the animals.

Chapter 5 investigates the population's conservation status. A deterministic population model was constructed, based on that of Zerbini et al. (2011), which describes the population's trajectory from the beginning of modern whaling in the area (i.e., 1900) to the present day. The model parameters and population trajectory are estimated using a Bayesian state-space framework, and using a different fitting algorithm (Markov chain Monte Carlo) of that used in Zerbini et al. (2011.). The inclusion of updated data on population growth rate (from the literature) and abundance estimates from Chapter 2 are also investigated. Conservation-relevant outputs include estimates of the maximum depletion and how close the population is to its historical carrying capacity.

In summary, previous work conducted in habitat use or distribution on humpback whales in Brazil are limited to descriptive analysis and restricted area (Martins et al., 2001; Zerbini et al. 2004), coarse resolution (Zerbini et al., 2004, Andriolo et al., 2010) and few environmental covariates investigated (Pavanato et al., 2018). The conservation status for this population was only directly investigated in Zerbini et al. (2011), although population dynamics parameters were also investigated in Freitas et al. (2004) and Wedekin et al. (2017). Therefore, updated and more complete assessment of population status is currently lacking. The present thesis aims to advance knowledge by quantitatively investigate habitat use and distribution considering a relatively wide range of environmental covariates, while producing robust abundance estimates; provide information on potential risks from oil extraction activities; and update information on population status.

Expected outcomes from this PhD thesis are a better understanding of the ecology of humpback whales wintering in Brazil. Through investigating their distribution, habitat use, potential impacts and conservation status, the implementation and/or need for management and conservation actions can be better informed. Important environmental features related to their habitat use are expected to be identified, as well as areas of concentration of animals and areas where animals are under higher risk of impact from oil spills.

CHAPTER 2

DISTRIBUTION AND ABUNDANCE OF WESTERN SOUTH ATLANTIC HUMPBACK WHALES FROM LINE TRANSECT DATA

This chapter has been published as:

Bortolotto, G.A., Danilewicz, D., Hammond, P.S., Thomas, L. & Zerbini, A.N. (2017) Whale distribution in a breeding area: spatial models of habitat use and abundance of western South Atlantic humpback whales. *Marine Ecology Progress Series*, 585:212-227.

ABSTRACT

The western South Atlantic humpback whale population was severely depleted by commercial whaling in the late 19th and 20th centuries, and today inhabits a human-impacted environment in its wintering grounds off the Brazilian coast. Whales are known to be distributed over the continental shelf during winter and spring, and to concentrate in the Abrolhos Bank, an enlargement of the shelf. Current information on habitat use is poor, which is important to inform conservation planning and management of human activities. In this chapter, distribution patterns and their relation to environmental features were investigated to explore habitat use and to generate distribution maps, and new estimates of population size were derived, which can be used in conservation status assessments. Spatial models were fitted to line transect data from 2 research cruises conducted in 2008 and 2012 to investigate (1) habitat use, (2) abundance and (3) distribution in relation to oil production fields for humpback whales wintering on the Brazilian continental shelf. Three models were fitted to address each of the points above. Candidate explanatory variables were year, depth, seabed slope, sea-surface temperature (SST), northing and easting, current speed, wind speed, distance to the coastline, to the continental shelf break, and to the nearest oil production field, and shelter (a combination of wind speed and SST categories). In the habitat use model (which excluded northing, easting and distance to oil production field *a priori*), whale density was estimated to be higher in slower currents, at shorter distances to both the coastline and the shelf break, at SSTs between 24 and 25°C and to be related to sheltered waters. In the abundance estimation model (which excluded only distance to oil production field *a priori*), easting

and northing were included in the model instead of SST. Estimated abundance was 14,264 whales (CV = 0.084) for 2008 and 20,389 (CV = 0.071) for 2012. Whale density was positively related to distances to oil fields at short distances, peaking at 100 km and negatively related to increasing greater distances. Environmental variables explained well the variation in whale density; higher density was found to the south of the Abrolhos Archipelago, and shelter were indicated to be important for these animals in their breeding area. Although no response to the presence of oil fields could be investigated here, whales can be expected to be found at not very short or very long distances to fields. Estimated distribution patterns presented here can be used to mitigate potential human-related impacts, such as supporting protection in the population's core habitat near the Abrolhos Archipelago.

2.1 INTRODUCTION

The Brazilian coast is inhabited every winter and spring by the western South Atlantic (WSA) humpback whale population (also referred to as breeding stock A by the International Whaling Commission; IWC, 1998). Whales aggregate in coastal waters along the central and north-eastern coasts of Brazil to mate and give birth before migrating south to feeding areas around the South Georgia and South Sandwich Islands, the Scotia Sea (Martins et al., 2001; Zerbini et al., 2006; Engel and Martin, 2009). This population was severely exploited by whaling between the late 19th and mid-20th centuries (Zerbini et al., 2011; Morais et al., 2017) to the point of near extinction in the 1950s but has since been recovering (Chapter 5; Zerbini et al., 2011, Wedekin et al., 2017). Today, the International Union for Conservation of Nature and Natural Resources (IUCN) lists the conservation status of this species as "Least Concern" (Cooke, 2018). Recent abundance estimates from ship-based line transect surveys suggest that the WSA population size was near 20,000 individuals in 2012 (Bortolotto et al., 2016a). However, that estimate was not computed for the entire area currently recognized as the typical distribution range of these animals during the breeding season. This increasing population currently faces an environment modified by human activities, including marine traffic (Bezamat et al., 2015), fishing (Rocha-Campos et al., 2011; Moura et al., 2013; Ott et al., 2016), coastal water pollution (Moura et al., 2013; Ott et al., 2016), noise pollution (Rossi-Santos, 2015) and activities related to the oil industry (Chapter 4; Iversen et al.,

2009; Martins et al., 2013; Ronconi et al., 2015; Rossi-Santos, 2015). Specifically, there is an increasing interest in oil and gas production activities in the area; according to the Brazilian National Agency of Petroleum, Natural Gas and Biofuels (*Agência Nacional do Petróleo, Gás Natural e Biocombustíveis*, ANP), the majority of Brazilian petroleum reserves is found in the marine environment (http://app.anp.gov.br). A method to investigate the potential impact from oil spills to whales in the area is presented in Chapter 5.

With human-related activities in the area increasing, negative interactions with humpback whales are likely to become more frequent (Andriolo et al., 2010; Martins et al., 2013). Existing marine protected areas (MPAs) alone provide very limited effective protection in the breeding grounds for this population, because they only cover a small fraction of the whales' range (Castro et al., 2014). Therefore, a broad understanding of their distribution patterns and habitat use is fundamental to inform management actions. Area-based management actions with the objective of protecting this population may additionally help protecting other marine species, given that whales occupy relatively large and biodiversity-rich marine habitat.

For seasonal migratory animals such as most baleen whale species, the environmental factors expected to be important in habitat selection differ between feeding areas, where prey distribution is the primary driver (e.g., Macleod et al., 2004; Friedlaender et al., 2006), and breeding areas (Corkeron and Connor, 1999). During the breeding season, large whales select habitat according to their breeding status (Rayment et al., 2015), presence of calves in groups (Cartwright et al., 2012) and other reproduction-related characteristics (Ersts and Rosenbaum, 2003; Craig et al., 2014; Lindsay et al., 2016). In this context, sheltered waters, bathymetric features, distance to the shore and sea-surface temperature (SST) have been indicated as important factors for habitat usage of humpback whales in breeding areas (e.g., Taber and Thomas, 1982; Smultea, 1994; Rasmussen et al., 2007; Félix and Botero-Acosta, 2011; Cartwright et al., 2012; Trudelle et al., 2016). Understanding and explaining key features of the ecology of migratory whale populations, such as habitat use, distribution and abundance, may provide important information for evaluating the impacts of human use of the environment inhabited by them.

WSA humpback whales are found in their breeding area, the Brazilian continental shelf between Natal (5° S) and Cabo Frio (23° S; Figure 2.1), during winter and spring every year, and animals concentrate on the Abrolhos Bank (~18° S) (Zerbini et al., 2006; Andriolo et al., 2010). The few previous studies that formally investigated their distribution relative to environmental variables (Wedekin, 2011; Pavanato et al., 2017), or how they use the available habitat (Martins et al., 2001), indicate that bathymetric features (e.g., slope, depth) may play an important role in how WSA whale groups are distributed.

This chapter provides new insights into the distribution and density of WSA humpback whales in relation to environmental features in their breeding grounds, and new abundance estimates for this population are also presented. Spatial models were applied to line transect data (i.e., density surface models, DSMs; Miller et al., 2013) from ship-based surveys conducted in 2008 and 2012 (Bortolotto et al., 2016a). Spatial models were fitted focussing on three main objectives: (1) to investigate habitat use, (2) to calculate model-based abundance estimates and (3) to investigate the distribution of whales in relation to oil production fields. The new information should inform management actions to conserve humpback whales on their Brazilian breeding grounds. More specifically, new abundance estimates may be used to update this population's conservation status (Chapter 5), and the distribution results to evaluate areas where this population may be at higher risk of being affected by human-related activities (Chapter 4).

2.2 MATERIALS AND METHODS

Shipboard visual line transect surveys were conducted in 2008 and 2012 during research cruises aboard the RV *Atlântico Sul* (*Universidade Federal do Rio Grande*, FURG). Cruises were part of the Monitoring Whales by Satellite Project (*Projeto Monitoramento de Baleias por Satélite*, PMBS). The main objectives of PMBS were to deploy satellite-linked tags on humpback whales to track their movements, to understand their space-use patterns in breeding and feeding grounds and to characterize their migratory routes (Chapter 3; Zerbini et al., 2006).

The survey area corresponded to the Brazilian continental shelf, between the shore and the shelf break (defined here as up to the 500 m isobath) from Cabo de São Roque (5° S),

in Rio Grande do Norte State, to Cabo Frio (23° S), in Rio de Janeiro State (Figure 2.1). Surveys were conducted from 25 August to 23 September in 2008 and from 7 August to 3 September in 2012, during the expected annual peak of occurrence of humpback whales in the area (August−September; Martins et al., 2001; Morete et al., 2003). Transect lines were designed to survey the full extent of this population's breeding area, and data collection followed conventional (i.e., "single-platform") distance sampling methodology (Buckland et al., 2001). Lines were surveyed in passing mode, when the survey platform does not approach the detected animals (Hammond, 2010). Survey design, observation effort and data collection details are detailed in previous work (Bortolotto et al., 2016a; b).

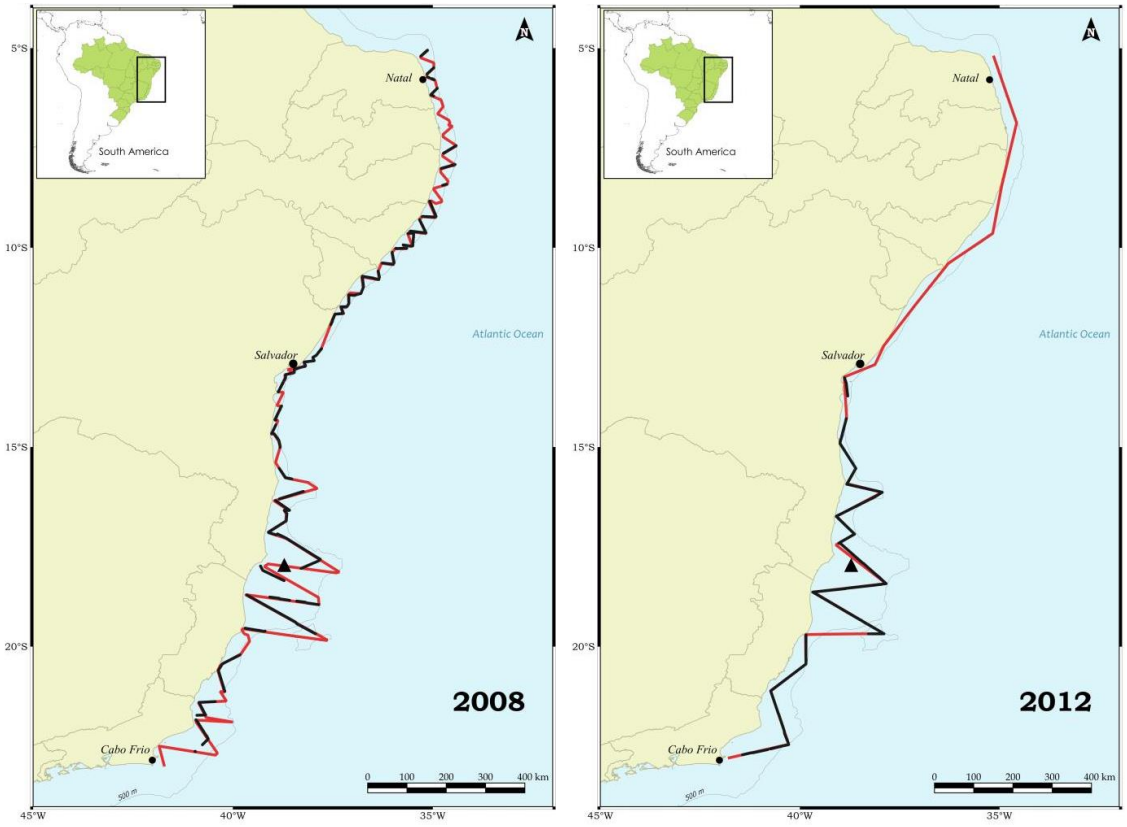


Figure 2.1. Survey lines in 2008 and 2012. Planned (red lines) and completed effort (black thick lines) are shown. Black triangles indicate the location of the Abrolhos Archipelago.

2.2.1 Correcting for imperfect detection: detection function modelling

In the detection function modelling, because other large whale species where rarely seen during the survey (1% of total sightings), sightings that were attributed to "unidentified large whales" (21% of total) were pooled with those of confirmed humpback whales. It

is very unlikely that unidentified whale sightings were not of humpback whales, as discussed in Bortolotto et al. (2016a).

Detection functions were fitted to perpendicular distance data using R (version 3.2.1; R Core Team, 2015) and the 'Distance' package (version 0.9.6; Miller, 2016). Key functions considered were the half-normal and the hazard rate (Buckland et al., 2001), with no adjustment terms. The truncation distance, i.e., the maximum distance between the trackline and sightings considered, was chosen by fitting preliminary detection function models to entire sightings dataset and identifying distances at which sightings presented very small detection probabilities (less than 0.15; Buckland et al., 2001). Factor covariates sea condition ("calm" for Beaufort 0 3 and "moderate" for Beaufort 4-6), detection cue (splash, body, blow or "other"), detection method (binoculars or naked eye) and year (2008 or 2012), and the numerical covariate group size (from 1 to 7) were considered. Variance in the detection function parameters was estimated using Fisher's information matrix (Buckland et al., 2001). Detection function model selection was performed by observing AIC scores in a forward-step approach: models with no covariates were fitted with both key functions, and the one presenting the smallest AIC was selected in that step. A set of new models was then created by separately including each one of the above covariates to the model selected in the previous step. If any of the models that included one covariate presented a smaller AIC when compared to the model selected in the previous step, the model with the smallest AIC was selected in the current step. This procedure was repeated until the inclusion of a new covariate did not cause reduction in AIC, and the final selected model was that which presented the smallest AIC among all that were fitted.

2.2.2 Data for spatial modelling

Survey transect lines were divided into 8 km segments using QGIS software (version 2.8.3; QGIS Development Team, 2015). Standard segment length was chosen to be twice the truncation distance (= 4 km), resulting in 8 by 8 km (64 km²) squares for most segments. During line segmentation, some segments at the end of lines were shorter than 8 km. In those cases, segments less than 4 km long were merged with the previous one and those longer than 4 km were considered as an independent new segment. A few segments (5 out of 516) that were less than 4 km long, and that could not be merged with

another line, were excluded from the analysis (including any observations made on those segments; in the event there were none). The response variable used to model whale distribution was the estimated abundance per segment as suggested in *dsm* R package documentation, with an offset for segment area.

Based on previous studies of the distribution of cetaceans in breeding areas, and also on environmental data availability, covariates considered as candidate explanatory variables were: current speed close to the surface, depth, distance to coast, distance to shelf break, distance to the nearest oil field, seabed slope, SST, wind speed at the surface, geographic position (northing and easting) and year (Table 2.1). Additionally, to represent a combination of environmental conditions that may be related to energy saving for calves, six categories for shelter (Table 2.1) were created by combining three categories of wind speeds at the surface ("light" for values between 0.94 and 5.15 m s^{-1}; "moderate" for values between 5.15 and 6.67 m s^{-1}; "strong" for values between 6.67 and 9.16 m s^{-1}) and two categories of SST ("cold" for values between the minimum of 20.2 and 24.7°C; "warm" for values between 24.7°C and the maximum 26.9°C). The wind and SST categories were delimited by quantiles of wind speed (33^{rd} percentile = 5.15 m s^{-1} and 66^{th} percentile = 6.67 m s^{-1}) and SST (median = 24.7°C).

Table 2.1. Explanatory variables tested in generalized additive models to model the density of humpback whales off the coast of Brazil. Resolution is given as spatial and/or temporal, depending on the covariate nature.

Variables	Description	Resolution	Unit	Reference/Data source
Curr.sp	Speed of the water current close to the surface	5 d; 0.33 × 0.33° (latitude × longitude)	m s^{-1}	OSCAR dataset (ESR, 2009)
Depth	Depth	0.1 × 0.1° (latitude × longitude)	m	ETOPO1 (Amante and Eakins, 2009)
Dist.coast	Distance to the coastline	–	m	SisCom (IBAMA, 2011)
Dist.shelf	Distance to the 500 m isobath	–	m	500 m isobath created from ETOPO1 in GIS software
Shelter	Category according to values of wind.sp and SST	–	–	–
Slope	Seabed slope: percentage of elevation over distance	0.1 × 0.1° (latitude × longitude)		Derived from ETOPO1
SST	Sea surface temperature	1 d; 0.011 × 0.011° (latitude × longitude)	°C	JPL-L4UHfnd-GLOB-MUR dataset (JPL MUR MEaSUREs Project, 2010)
Oil.field	Distance to the nearest oil production field	–	m	ANP (ANP, 2017)
Wind.sp	Speed of wind at the surface	6 h (the daily mean was used); 80 × 80 km	m s^{-1}	ERA-Interim dataset (Dee et al., 2011)
x	Easting	–	m	Survey GPS
y	Northing	–	m	Survey GPS
Year	Year of survey	–	yr	Survey data

Values for depth were extracted from the global model of land topography and ocean bathymetry ETOPO1 (Amante and Eakins, 2009). Circular buffers (radius = 4 km) were created around segment midpoints in QGIS, and the average of depth values within the buffer zone was computed for each segment. This procedure was adopted because the resolution of ETOPO1 was much finer than the size of segments and buffers (between 13 and 16 ETOPO1 cells were included in the 50 km^2 buffers and used to compute mean

depth values). After extraction of mean depth values, 25 out of 511 segments gave values greater than 500 m and were excluded from the analysis (together with any sightings they contained) because the study area was previously defined as the continental shelf, from the shore up to the 500 m isobath. Slope values were derived from ETOPO1 data and were obtained in the same way, i.e., extracting mean values using the same circular buffers.

Distances to physical features (distance to coast, distance to shelf break and distance to oil fields) were calculated in QGIS or R as the shortest distance between the segment midpoint and the feature. For the distance to coast variable, the Brazilian coastline was obtained from a shapefile provided by SisCom (IBAMA, 2011). To represent the continental shelf break, the 500 m isobath was generated from ETOPO1 in ArcGIS software using the "contour tool" function (ArcGIS Desktop: release 10, ESRI). Locations of oil fields at the production stage during the time of the surveys in Brazilian coastal waters were obtained from the website of the National Agency of Petroleum, Natural Gas and Biofuels (ANP; ANP, 2017). Polygons of oil production fields from the marine portions (Figure 2.2) of the six sedimentary basins covered in the survey area (Alagoas, Camamú, Campos, Espírito Santo, Recôncavo and Sergipe) were extracted from an online map available from the ANP website.

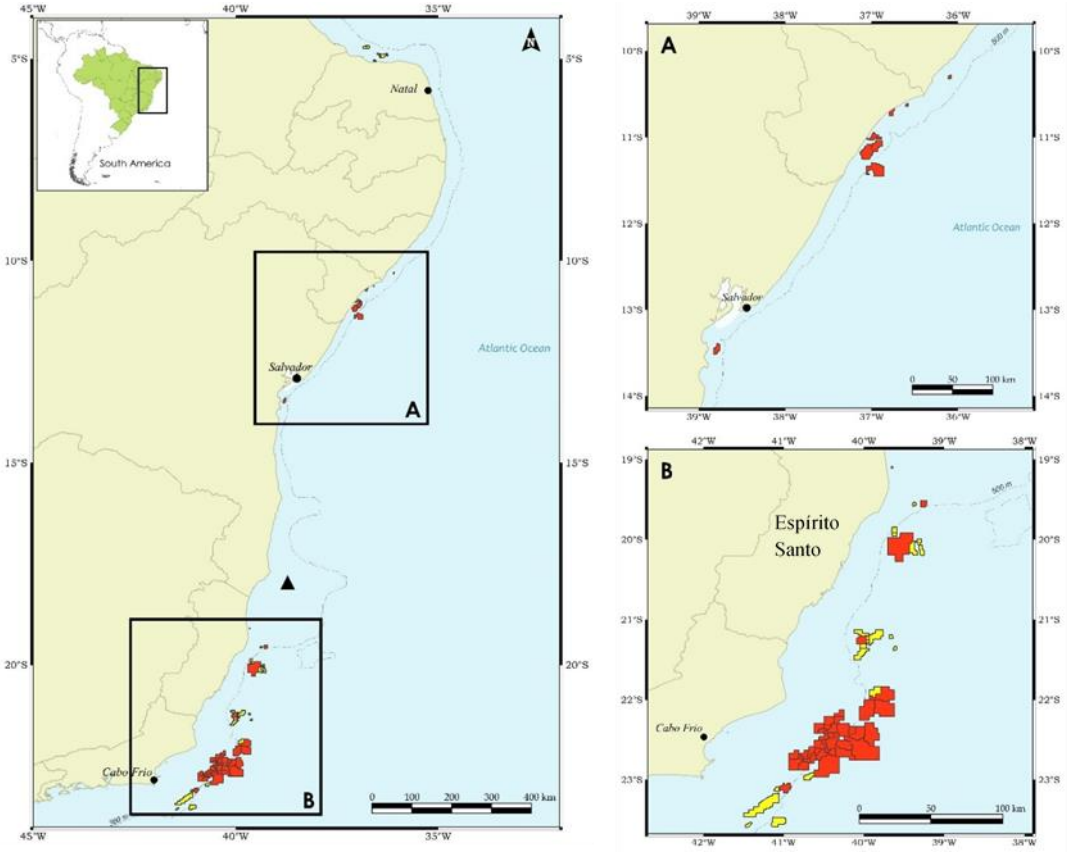


Figure 2.2. Oil production fields' locations. Red polygons represent fields that were present since 2008. Yellow polygons represent additional fields in 2012.

SST was extracted from the "MUR Global Foundation Sea Surface Temperature Analysis" dataset (JPL MUR MEaSUREs Project, 2010) and ocean currents from the 'OSCAR' dataset (ESR, 2009). Wind speed data were extracted from the 'ERA-Interim' dataset (Dee et al., 2011). With the exception of SST, the resolution of these datasets was too coarse when compared to the size of the circular buffers, so segment midpoints were used to extract covariate values in R software ('raster' package; Hijmans, 2016). For SST, the circular buffers previously described were used to obtain mean values (around 40 SST values per buffer).

2.2.3 Spatial models and model selection

An initial investigation was performed to assess correlation among explanatory variables, and those that were highly correlated (i.e., a pair of variables with Pearson's correlation coefficient greater than 0.7) were not included in the same model at the same time. Interaction terms, combining year and other covariates, were not tested because part of

the study area was not surveyed in 2012, which would make the comparison severely unbalanced.

The quasi-Poisson distribution (commonly adequate for over-dispersed count data) with a logarithmic link function was assumed for the response variable; negative binomial and Tweedie distributions were also tested, and model diagnostic plots such as the "Normal Q-Q-plot" (gam.check R function, mgcv package), supported this decision. Generalized additive models (GAMs) were fitted using the dsm R package (version 2.2.14; Miller et al., 2017). Smooth functions were fitted to covariates, with a bivariate smooth for geographic position (easting and northing). The basis dimension parameter *k* for the geographic position smooth term was set to 20, and for the univariate smooth terms it was set to eight (see Wood [2006] for guidance on setting the dimension parameter). Model selection was conducted using a forward approach (i.e., adding one variable at a time), starting with a set of models, each with only a single candidate explanatory variable. The model selected at each step was chosen by looking for an improvement in the restricted maximum likelihood (REML; Harville, 1977) score. This score was used to bypass problems with parameter estimation that other potential scores (e.g., UBRE and GCV) may present when applying DSMs, following recommendations from Miller et al. (2013). However, models were also fit using GCV and the GCV score was assessed to check if it supported the inclusion or dropping of covariates, as indicated by the REML score. Auto-correlation in the residuals (ordered by the time of data collection, for each survey) of spatial models was checked using the acf function (stats R package; R Core Team, 2015). Model adequacy was assessed with model diagnostic plots (function gam.check, mgcv R package).

Three modelling exercises were undertaken, each considering a different set of covariates and having different objectives:

1. Habitat use model (HUM): to explain habitat use in a way that could be interpreted biologically; all variables, except distance to oil fields and geographic position (northing/easting), were considered;
2. Abundance estimation model (AEM): to compute abundance estimates from the spatial model; all available variables were considered;

3. Oil fields model (OFM): to investigate the distribution of whales in relation to oil production fields; the covariate representing the distance to the nearest oil field was included in the model selected in the HUM described above.

The HUM was designed to investigate which of the available environmental variables best explained variation in whale density, while the AEM was designed to obtain the best density surface prediction, possibly including northing/easting, which could explain variability that was not explained by the other environmental covariates. The OFM exercise aimed to evaluate how much variability in whale density could be attributed to distance from oil fields.

2.2.4 Predictions

A prediction grid formed by 8 × 8 km cells was created over the entire study area using QGIS. The size of the prediction grid cells was chosen to match that of the segments used in the models. Covariate values for each grid cell were obtained in a similar way to that described for segments, using cell midpoints or buffers around midpoints. For covariates that varied in time within each survey (e.g., SST), the mean of values for the survey period was used for predictions. No predictions were derived from the OFM, because objectives for this model were to investigate the relationship between distance to oil fields and whale density, which are illustrated with fitted smooth function plots, without the need for maps.

Model-based abundance estimates for 2008 and 2012 were obtained from sums across all grid cells of predicted values from the AEM, for each year. Maps showing patterns of distribution (density surface) were created using the AEM predictions in QGIS. Variances were obtained with the delta method, combining the variance from the detection function and the spatial models, using the "dsm.var" function of the "dsm" R package. AEM maps of uncertainty in model predictions (coefficient of variation surface) were also created, by dividing the square root of the variance by the point estimate on each grid cell. Predictions in 2012 were extrapolated to the area to the north of Salvador (~13°S), which was not surveyed in that year (Figure 2.1) because of poor weather conditions (Bortolotto et al., 2016a).

2.3 RESULTS

Survey effort used in the analysis totalled 2,350 km in 2008 and 1,700 km in 2012. The number of whale groups (including all whale group categories, such as mother-calf pairs, and solitary animals) in the data was 493 (416 humpbacks and 77 unidentified large whales) and 737 (557 humpbacks and 180 unidentified large whales) in 2008 and 2012, respectively. Paired covariate plots and summary of correlation are presented in Figure 2.3.

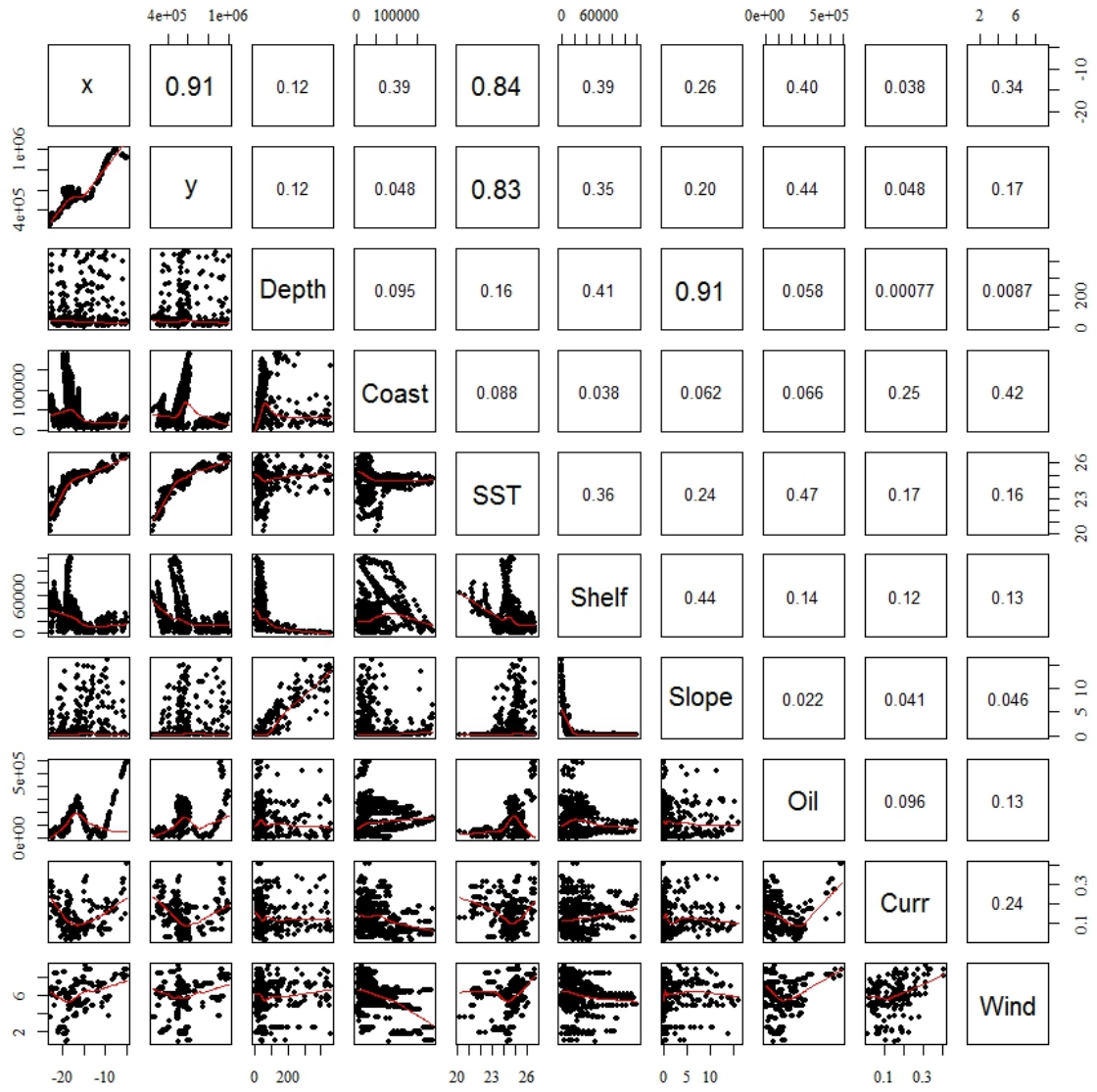


Figure 2.3. Paired covariate plots (lower left diagonal) and correlation (upper right diagonal). Correlation font size is larger when correlation is high. (Curr = current speed, Wind = wind speed, Shelf = shelf distance, coast = coast distance).

2.3.1 Detection function

Perpendicular distances were truncated at 4 km, resulting in 81 (out of 1,230) detections being excluded from the detection function analysis. The best-fitting detection function based on the AIC score was a hazard rate model with the covariates cue, year and sea conditions (Figure 2.4; Table 2.2). The average probability of detection *p* was estimated as 0.482 (CV = 0.044) and the goodness of fit tests showed a good fit (Kolmogorov-Smirnov test statistic = 0.016, p = 0.930; Cramer-von Mises test [unweighted] statistic = 0.036, p = 0.952).

Table 2.2. Detection function parameters from a hazard-rate key-model fitted to 1,149 perpendicular distance values for humpback whale sightings (data were truncated at 4 km). Coefficient values are on the scale of the log link function. The intercept includes terms "cue blow", "year 2008" and "sea state calm".

Scale Coefficients	**Estimate**	**Standard error**
Intercept	7.097	0.125
Cue splash	0.535	0.162
Cue body	-0.470	0.164
Cue "other"	0.363	0.310
Year 2012	0.291	0.107
Sea state moderate	-0.220	0.107

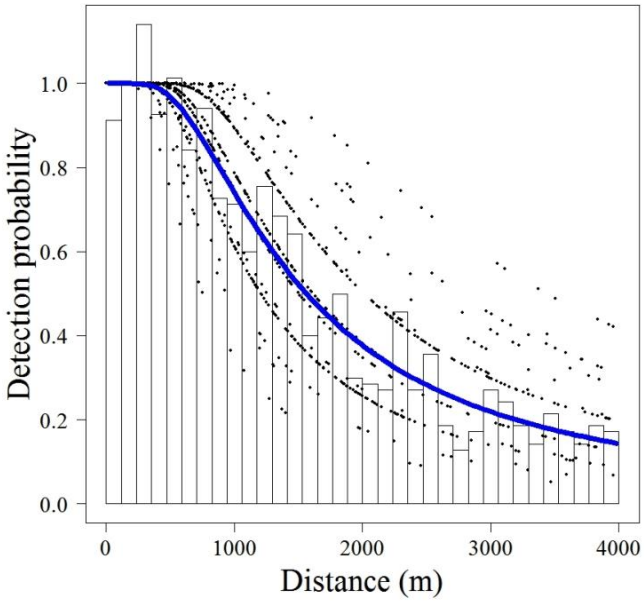


Figure 2.4. Detection function curve (blue line) from a hazard rate model fitted to the perpendicular distances (in metres) of humpback whale groups detected. Different dotted curves represent different combinations of the covariates sea condition, cue and year. Each point represents the predicted value for a single observation.

2.3.2 Spatial models

Model diagnostics (Figure 2.5, Figure 2.6 and Figure 2.7) indicated the quasi-Poisson distribution to be adequate. SST was highly correlated with geographic position. Depth, slope and distance to the shelf break were also correlated with one other. Therefore, if one of the above variables was selected at a model selection step, those correlated with it were not considered in subsequent steps of model selection. GCV scores supported the inclusion/exclusion of covariates as indicated by the REML score.

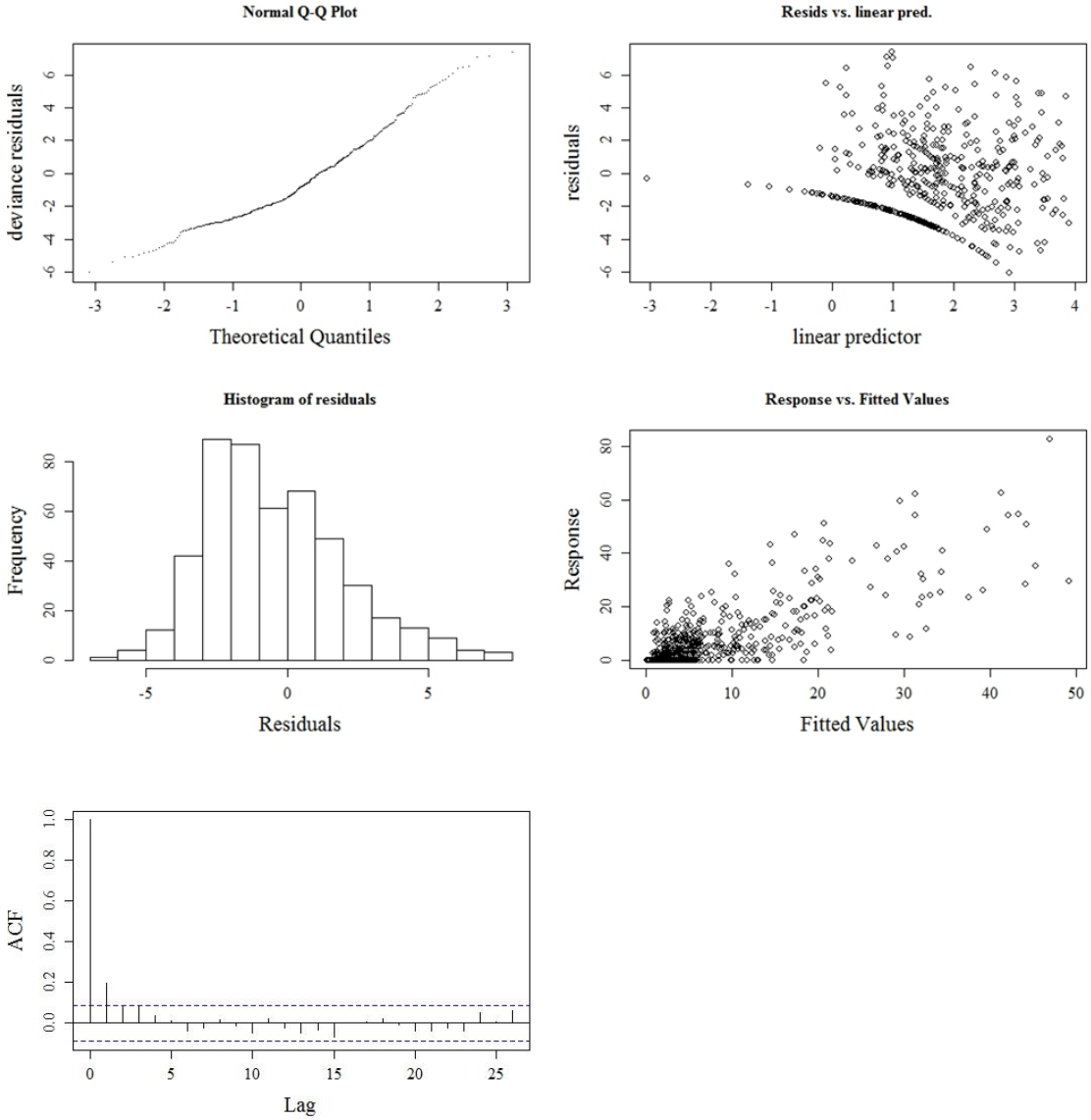


Figure 2.5. Habitat use model (HUM) diagnostic plots from "gam.check" R function and auto-correlation regression plot from "acf" function.

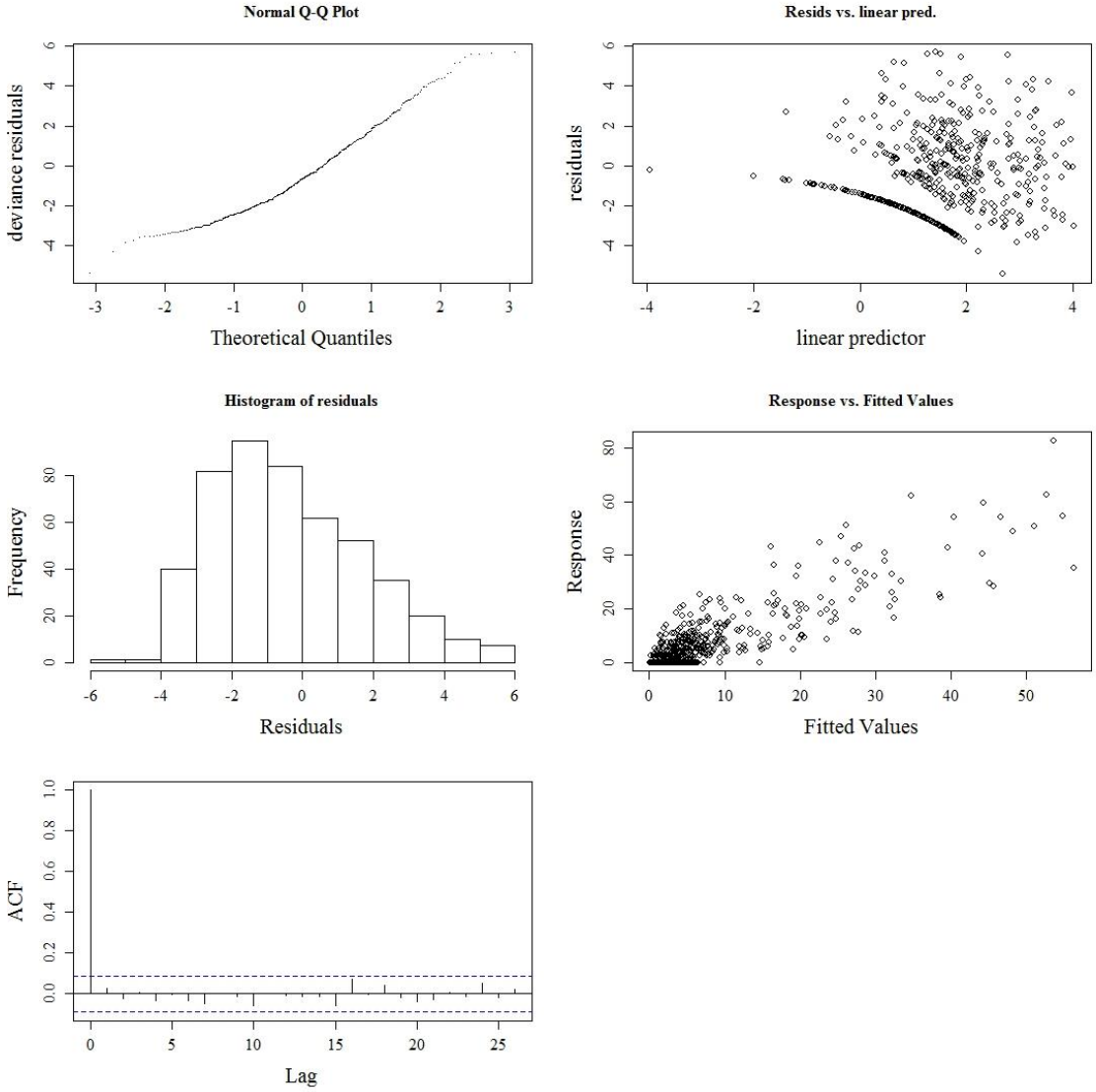


Figure 2.6. Abundance estimation model (AEM) diagnostic plots from "gam.check" R function and auto-correlation regression plot from "acf" function.

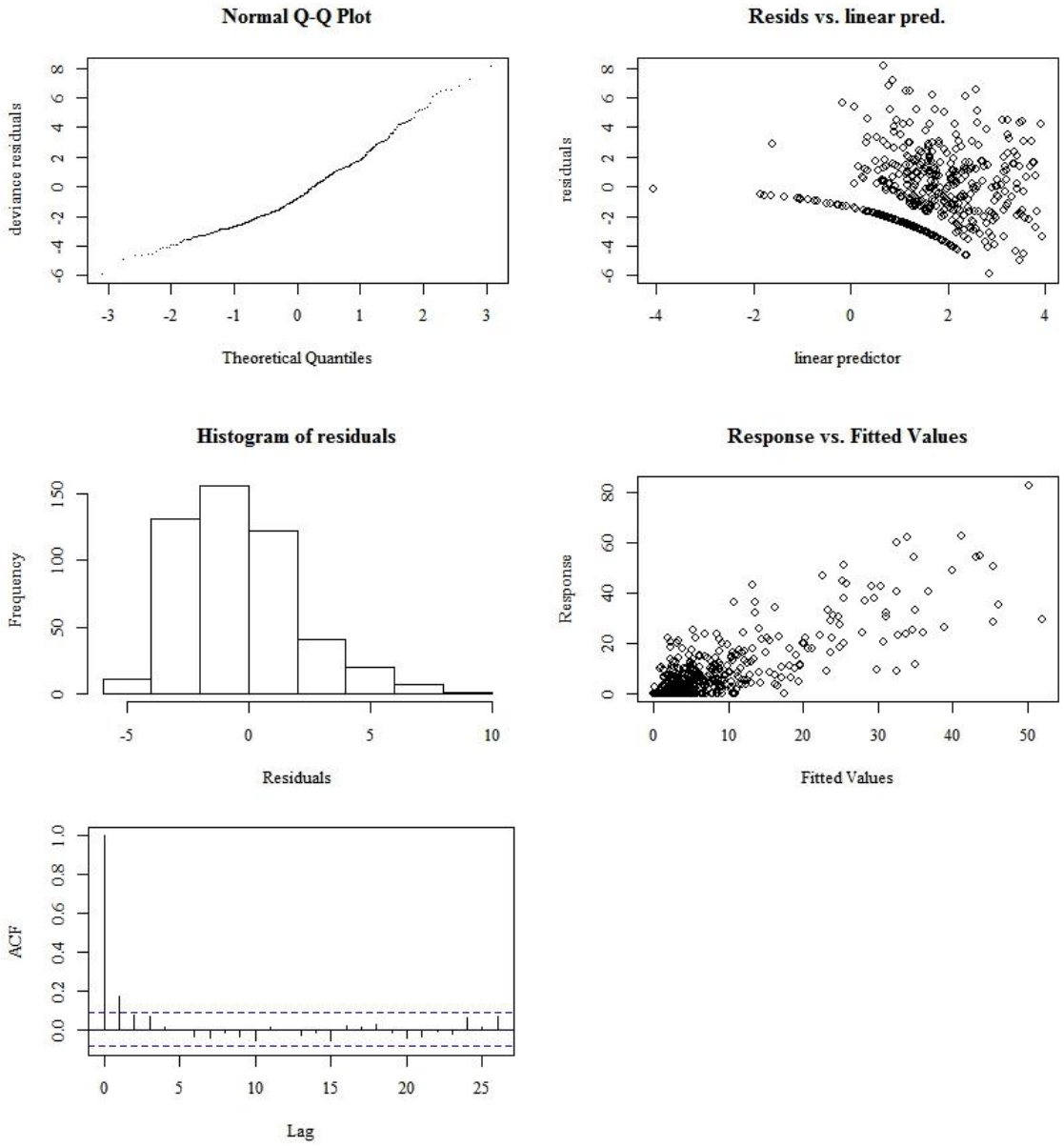


Figure 2.7. Oil fields model (OFM) diagnostic plots from "gam.check" R function and auto-correlation regression plot from "acf" function.

The final HUM included the variables distance to coast, distance to shelf break, SST, current speed and shelter, and explained 54.1% of the deviance. The variable with the most pronounced effect was SST, with a peak around 24−25°C (Figure 2.8). Whale density was positively related to distance to coast and distance to shelf break. It presented a gradual positive relationship to current speeds up to 0.2 m s^{-1}, peaking at that value, and was negatively related to greater current speeds. Shelter coefficients indicated differences

in whale densities between shelter categories, with significantly (at $\alpha = 0.05$) higher densities in relatively cold waters with light winds (Table 2.3; Figure 2.8).

Table 2.3. Parametric coefficients and smooth terms in the habitat use model (HUM). (t = t distribution value, df = degrees of freedom, F = F distribution value). *Significant at $\alpha = 0.05$

Coefficients	**Estimate**	**Standard error**	**t**	**p-value**
Intercept	-15.704	0.116	-134.819	< 0.001*
shelter.cold.moderate	-0.473	0.111	-4.272	< 0.001*
shelter.cold.strong	-1.122	0.271	-4.138	< 0.001*
shelter.warm.light	-0.760	0.193	-3.942	< 0.001*
shelter.warm.moderate	-1.140	0.261	-4.364	< 0.001*
shelter.warm.strong	-0.524	0.242	-2.164	0.031
Smooth terms	**Effective df**	**Reference df**	**F**	**p-value**
s(sst)	3.766	7	6.347	< 0.001*
s(dist.shelf)	0.975	7	5.041	< 0.001*
s(coast)	2.401	7	4.918	< 0.001*
s(curr.sp)	3.315	7	2.535	< 0.001*

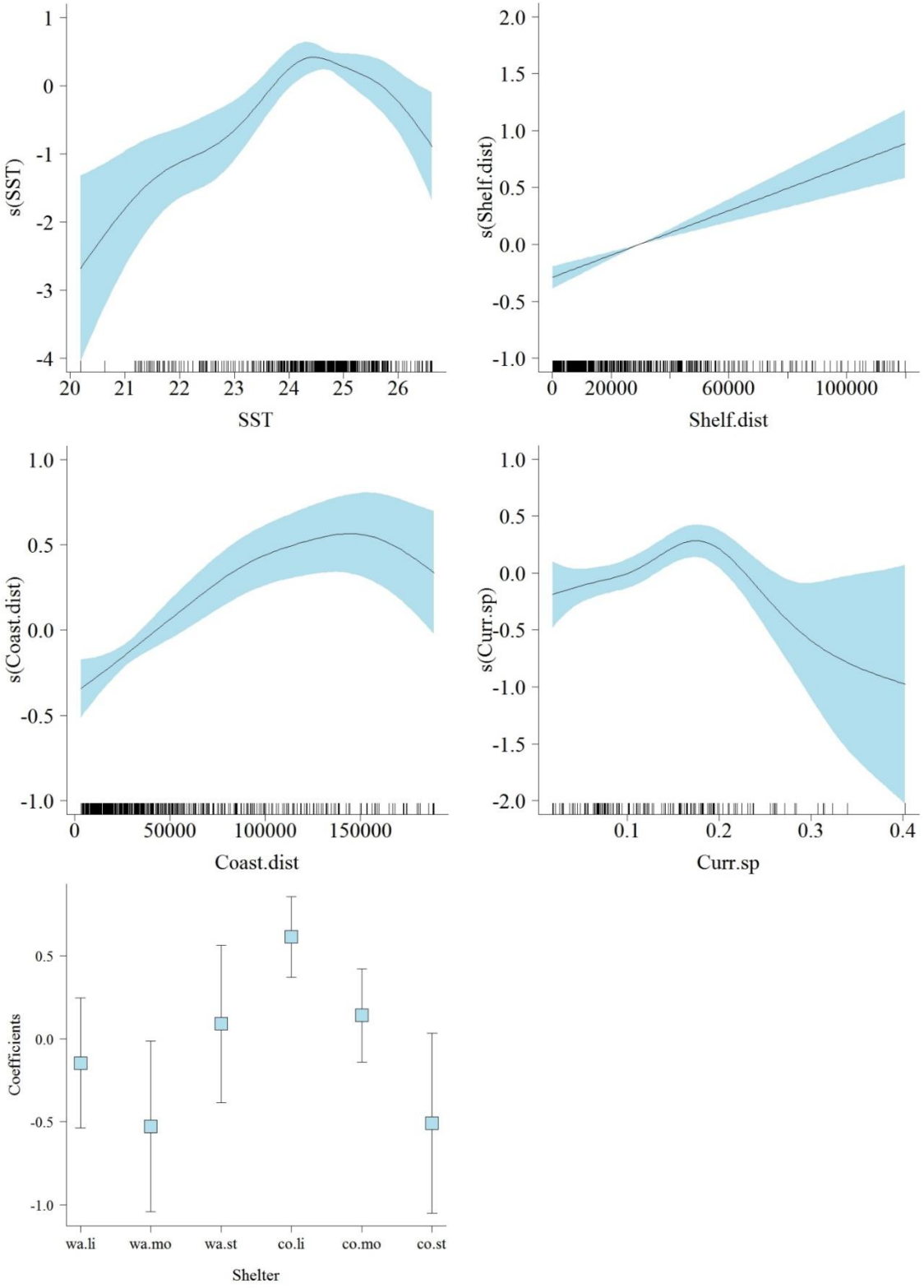


Figure 2.8. Fitted model terms for the habitat use model (HUM) of humpback whales Megaptera novaeangliae off the coast of Brazil. Shelter coefficients are presented relative to the intercept, i.e., "shelter.cold.light" in Table 2.3 (wa = warm SST, co = cold SST, li = light wind, mo = moderate wind, st = strong wind).

The selected AEM included the variables distance to coast, distance to shelf break, current speed, shelter and geographic position (Table 2.4; Figure 2.9), and had an explained deviance of 66.8%. Distance to coast presented an overall negative relationship to whale density, peaking at about 30 km, in contrast to the HUM for which that covariate presented an apparent peak at 150 km. Because this model explained a larger percentage of the deviance, the distribution patterns are better represented in the AEM map.

Table 2.4. Parametric coefficients and smooth terms in the abundance estimation model (AEM). (t = t distribution value, df = degrees of freedom, F = F distribution value). *Significant at $\alpha = 0.05$

Coefficients	**Estimate**	**Standard error**	**t**	**p-value**
Intercept	-16.007	0.105	-153.078	< 0.001*
shelter.cold.moderate	-0.279	0.109	-2.559	0.011*
shelter.cold.strong	-0.830	0.247	-3.364	< 0.001*
shelter.warm.light	-0.484	0.148	-3.268	0.001*
shelter.warm.moderate	-0.532	0.221	-2.402	0.012*
shelter.warm.strong	-0.470	0.207	-2.272	0.024*
Smooth terms	**Effective df**	**Reference df**	**F**	**p-value**
s(x,y)	15.865	19	9.911	< 0.001*
s(curr.sp)	3.294	7	4.009	< 0.001*
s(coast)	5.528	7	4.283	< 0.001*
s(dist.shelf)	0.940	7	2.155	< 0.001*

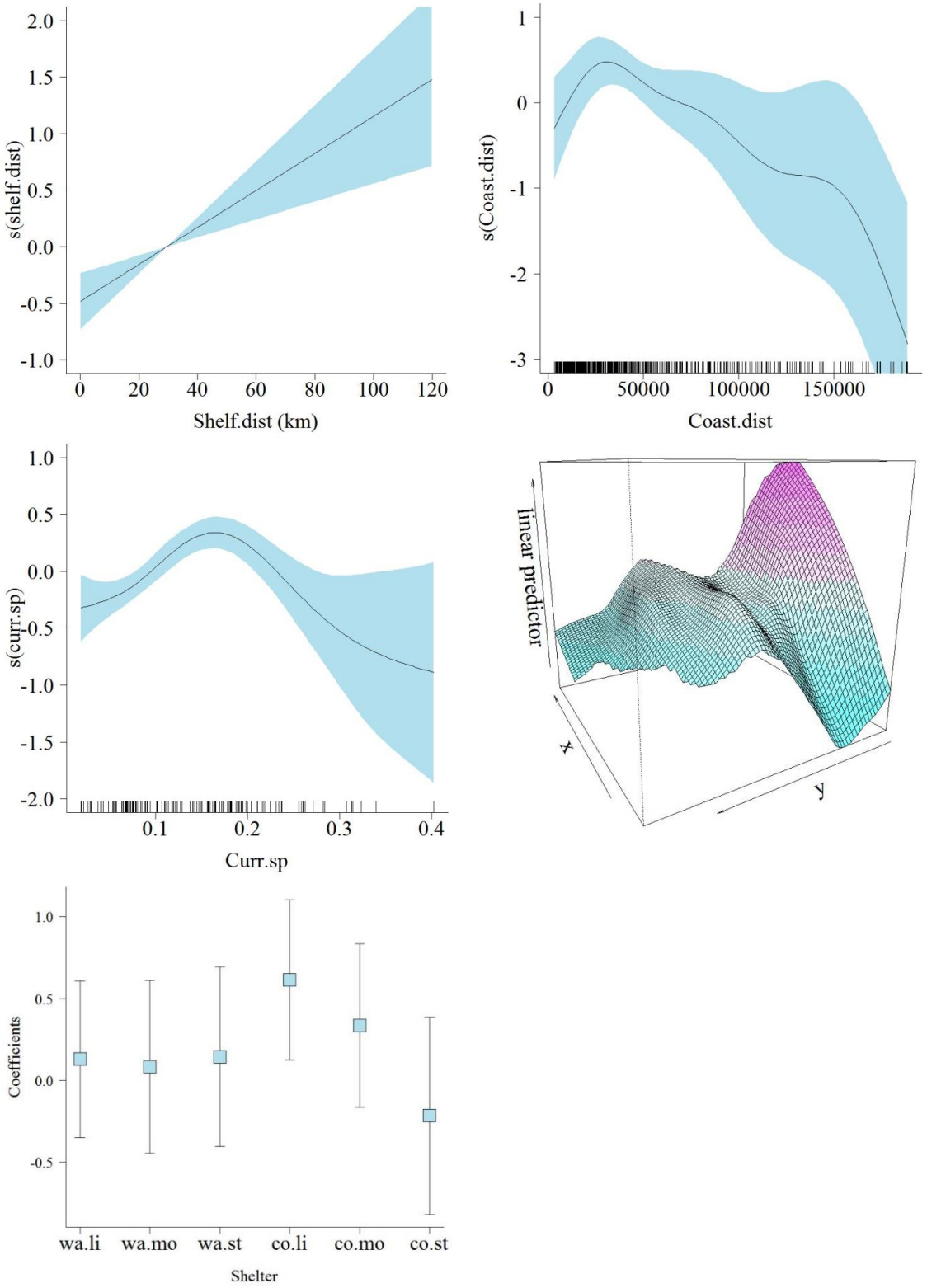


Figure 2.9. Fitted model terms for the abundance estimation model (AEM). Plot for x and y shows a contour of the model predictions, where all but easting and northing values are fixed (vis.gam function, mgcv R package, version 1.8; Wood 2011). Shelter coefficients are presented relative to the intercept, i.e.,

"shelter.cold.light" in Table 2.4 (wa = warm SST, co = cold SST, li = light wind, mo = moderate wind, st = strong wind).

The inclusion of the distance to oil production fields in the HUM, resulting in the OGM, caused little difference on in the relationship between the response and the other (already in the HUM) explanatory variables (compare Figure 2.8 and Figure 2.10). Although the new term was negatively related to density of whales, i.e., abundance decreased with increasing distance to oil fields, it showed a peak at around 100 kilometres (Figure 2.10). The percentage of explained deviance increased from 54.3 in the HUM to 57.2% in the OFM.

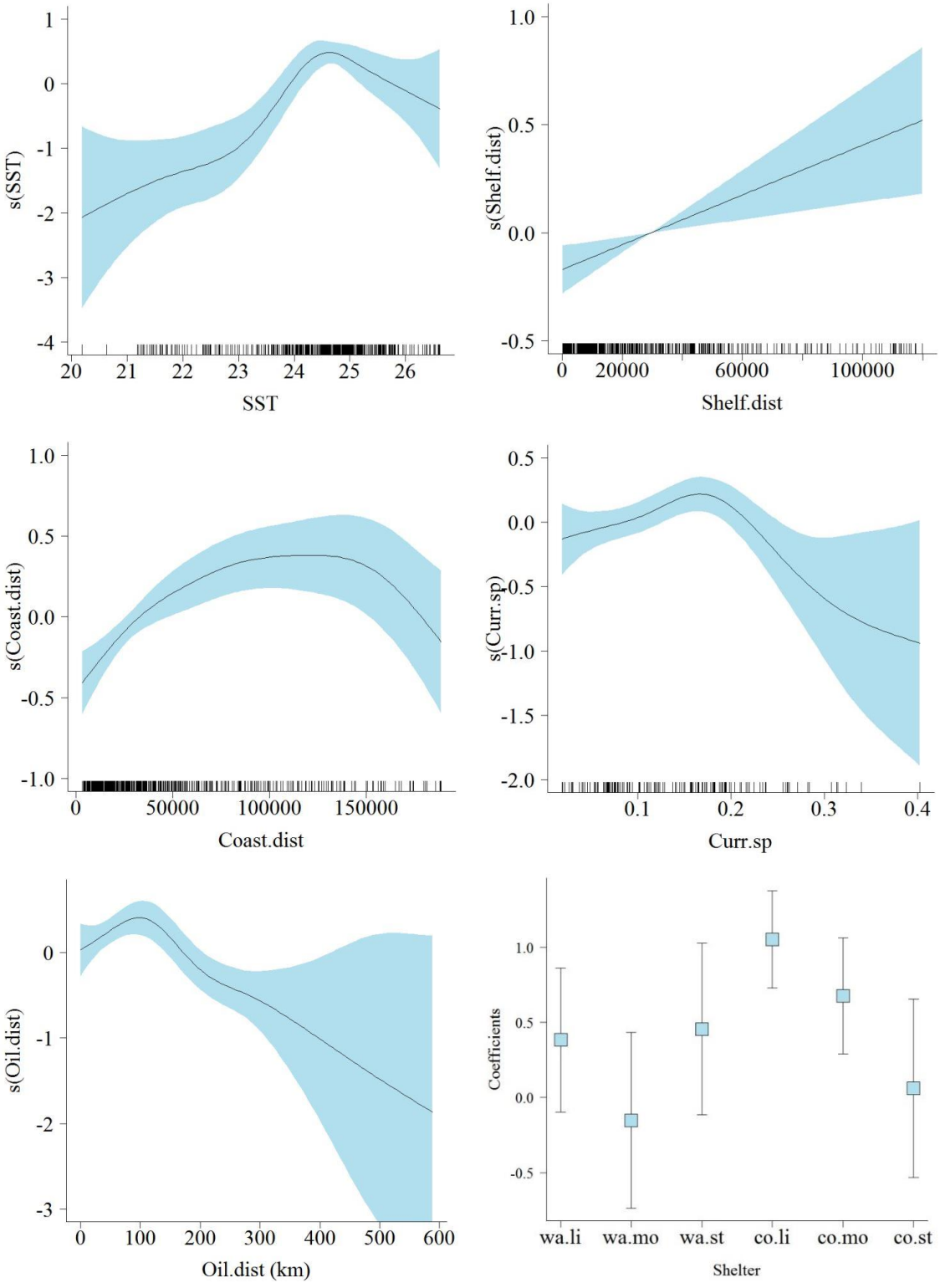


Figure 2.10. Fitted model terms for the oil fields model (OFM). Shelter coefficients are presented relative to the intercept, i.e., "shelter.cold.light" in Table 2.5. (wa = warm SST, co = cold SST, li = light wind, mo = moderate wind, st = strong wind).

Table 2.5. Parametric coefficients and smooth terms in the oil field model (OFM). (t = t distribution value, df = degrees of freedom, F = F distribution value). *Significant at $\alpha = 0.05$

Coefficients	**Estimate**	**SE**	**t**	**p-value**
Intercept	-15.755	0.1237	-127.346	< 0.001*
shelter.cold.moderate	-0.375	0.1141	-3.289	0.001*
shelter.cold.strong	-0.991	0.2692	-3.679	< 0.001*
shelter.warm.light	-0.670	0.1981	-3.380	< 0.001*
shelter.warm.moderate	-1.204	0.2775	-4.341	< 0.001*
shelter.warm.strong	-0.597	0.2554	-2.337	0.0199*
Smooth terms	**Effective df**	**Reference df**	**F**	**p-value**
s(SST)	3.773	7	7.222	< 0.001*
s(shelf.dist)	0.912	7	1.351	< 0.001*
s(coast.dist)	2.836	7	3.168	< 0.001*
s(curr.sp)	3.004	7	1.908	0.0016*
s(oil.dist)	3.488	7	3.985	< 0.001*

Very weak signs of auto-correlation were found in the residuals of HUM and OFM, and no signs of auto-correlation were present in the residuals of AEM (ACF plots; Figure 2.5, Figure 2.6 and Figure 2.7). A summary of covariates retained in the final three models is presented in Table 2.6.

Table 2.6. Generalized additive model results for the habitat use model (HUM), the abundance estimation model (AEM) and oil fields model (OFM). Variables are described in Table 2.1. Effective degrees of freedom for smooth terms (s) are presented inside brackets. Blank spaces represent variables not selected, and a dash represents a covariate not considered in the model selection. (REML = restricted maximum likelihood, F = factor).

Variable	HUM	AEM	OFM
curr.sp	s(3.315)	s(3.294)	s(3.004)
depth			
coast.dist	s(2.401)	s(5.528)	s(2.836)
shelf.dist	s(0.975)	s(0.940)	s(0.912)
shelter	*F*	*F*	*F*
slope			
SST	s(3.766)		s(3.773)
oil.dist	—	—	s(3.488)
wind.sp			
x, y	—	s(15.865)	—
year			
% Deviance explained	54.1	66.8	57.2

2.3.3 Abundance estimates

Estimated abundances for prediction grid cells ranged from 0.139 to 53.0 individuals (mean = 7.47, SD = 8.90) in 2008 and from 0.144 to 60.9 individuals mean = 10.7, SD = 12.7) in 2012., for grid cells sized 64 km². Model-based abundance estimates were 14,264 whales (CV = 0.084) for 2008 and 20,389 (CV = 0.071) for 2012 (Table 2.7). Surface maps for predicted density showed higher numbers in the Abrolhos Bank region, with a concentration of animals to the south of the Abrolhos Archipelago, which was more pronounced in 2012 (Figure 2.11). Uncertainty was higher in the north and south extremes of the survey area (Figure 2.12). Other areas also showed relatively high densities, such as the coast of Alagoas and Sergipe States (Figure 2.13), and near the city of Salvador, Bahia State (Figure 2.14).

Table 2.7. Summaries of uncertainty in the abundance estimation model (AEM) calculated analytically for GAM, with delta method, for 2008 and 2012.

2008		
Approximate asymptotic confidence interval		
2.5%	**Mean**	**97.5%**
12,108	14,264	16,805
Abundance		
Point estimate		14,264
CV of detection function		0.044
CV from GAM		0.071
Total standard error		1,195
Total coefficient of variation		0.084
2012		
Approximate asymptotic confidence interval		
2.5%	**Mean**	**97.5%**
17,746	20,389	23,426
Abundance		
Point estimate		20,389
CV of detection function		0.044
CV from GAM		0.056
Total standard error		1,446
Total coefficient of variation		0.071

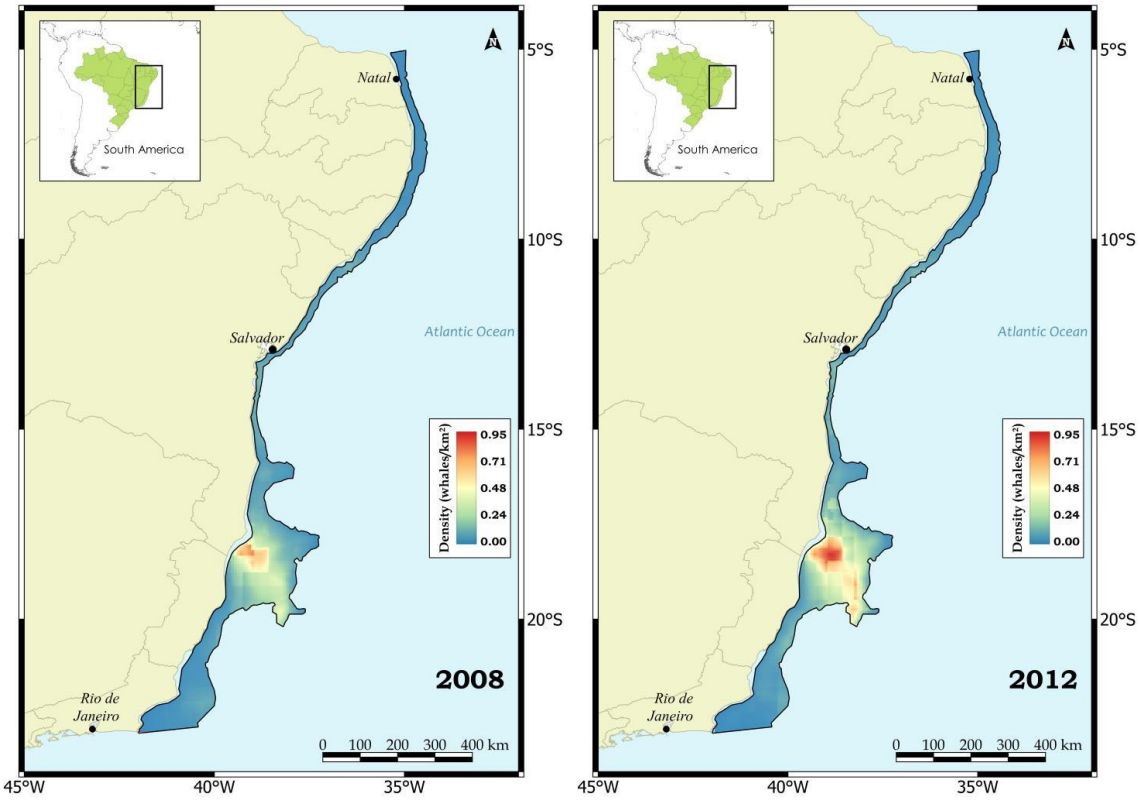


Figure 2.11. Density surface maps for 2008 and 2012. Predictions were made with the abundance estimation model (AEM).

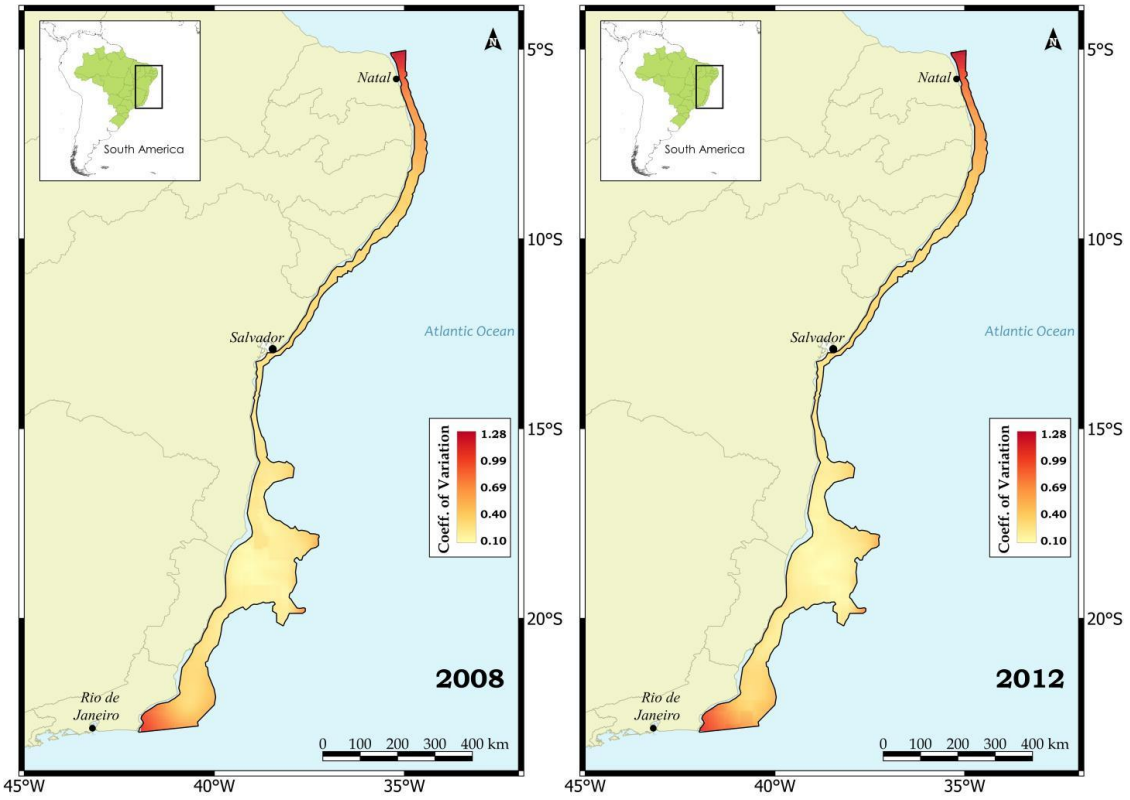


Figure 2.12. Coefficient of variation surface maps for 2008 and 2012 for the AEM.

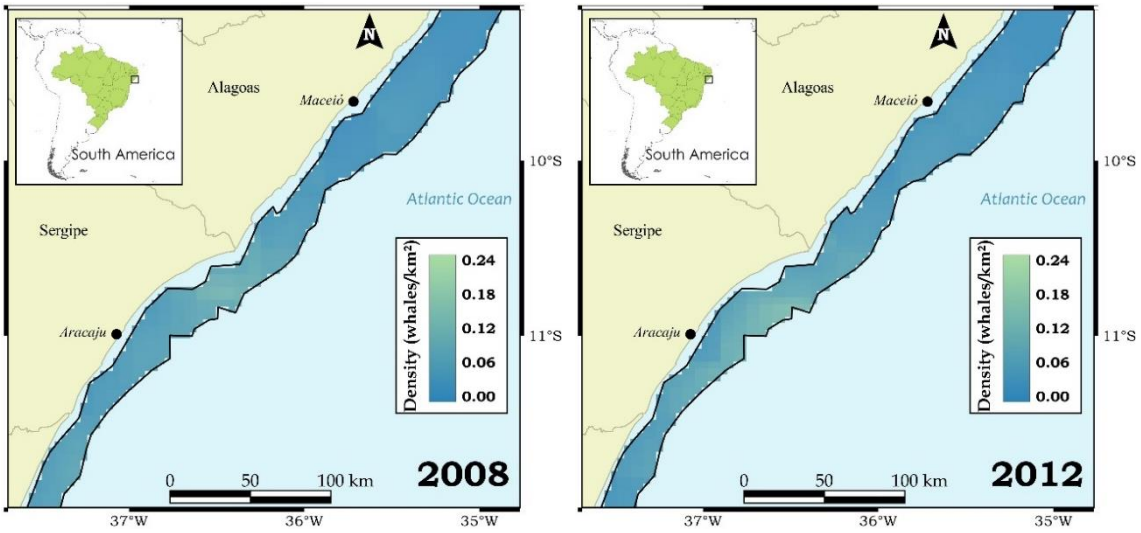


Figure 2.13. Density surface maps for 2008 and 2012 for the region of Sergipe and Alagoas coasts for the AEM.

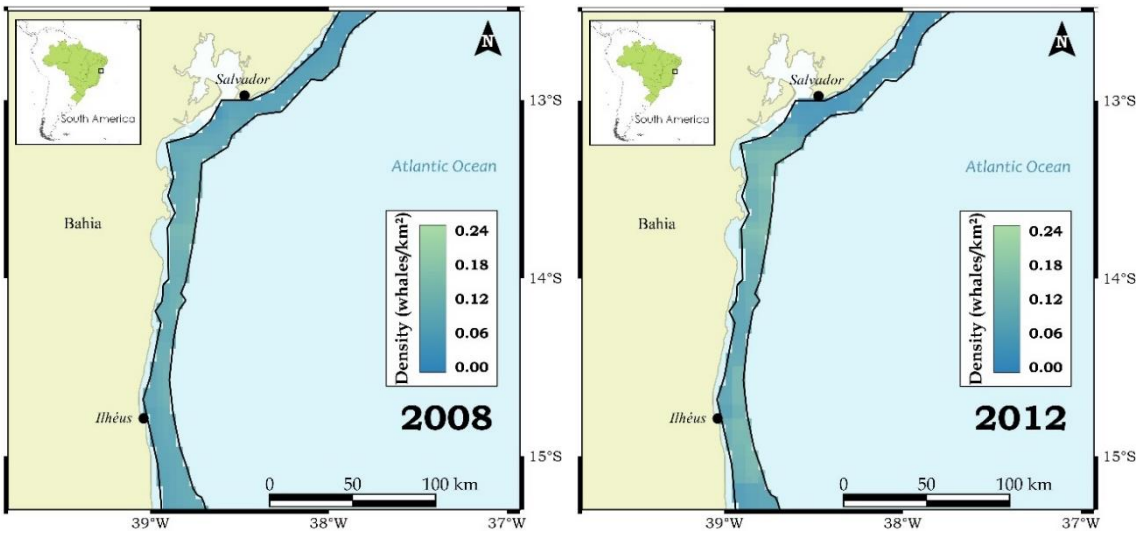


Figure 2.14. Density surface maps for 2008 and 2012 for part of the coast of Bahia State for the AEM.

2.4 DISCUSSION

Systematically collected sightings data were used to model the distribution and abundance of humpback whales in their wintering areas off the coast of Brazil. The suite of covariates tested included strong predictors of whale density across the study area, with SST and geographic position being the most powerful explanatory terms. The effect of year was not selected in the spatial models, suggesting that differences in the distribution patterns from 2008 to 2012 were better explained by the variation in the spatial covariates than by temporal changes between survey years.

These sightings data were previously used to estimate abundance of humpback whales off the coast of Brazil in 2008 and 2012 using design-based methods, such as distance sampling (Bortolotto et al., 2016a). However, the realized effort in that study did not completely reflect the designed transect lines. For example, because of unfavourable weather conditions in 2012, no data were available for areas to the north of Salvador, Bahia State (Figure 2.1). Consequently, the abundance estimate previously presented for that year was computed for only part of what is currently known to be the typical breeding area for WSA humpback whales.

Migratory whales show marked differences in habitat preferences according to different age classes, sexes, reproduction-related individual characteristics and/or group composition (Craig and Herman, 2000; Ersts and Rosenbaum, 2003; Elwen and Best, 2004a; Oviedo and Solís, 2008; Cartwright et al., 2012; Craig et al., 2014; Rayment et al., 2015), and for specific group types (Elwen and Best 2004b; Félix and Botero-Acosta, 2011) when in breeding areas. Therefore, information on the above is highly desirable for studies such as this. However, the passing mode data collection procedure adopted here prevented more specific data on individual whales, such as sex, age class or accurate group composition, from being obtained. Because of this, results presented here are representative of the population as a whole, not of any particular sex, age or group type. Although some of the results may be consistent with what could be expected for habitat preferences of breeding and/or calving animals in the area, such as the importance of shelter as a predictor of density, it is not possible to make robust inferences for specific reproductive stages. Another option to investigate habitat use and potentially include individual information is by using data from satellite tracking (Trudelle et al., 2016). Because the procedure of attaching tags requires close proximity to the animals, collection of individual and group information is possible at the moment of tagging. Although this is not the approach taken in Chapter 3 (which addresses analysis of tag data), investigations of habitat use of WSA humpback whales, relating whale movements and behaviour states to environmental features, are currently underway, which are expected to provide information on predictors of distribution and habitat use in relation to sex and group from the perspective of individuals and/or considering behavioural states.

2.4.1 Spatial modelling

Covariates retained in the final models explained high percentages of the variation in whale density across the surveyed area (deviance explained = 54.1% for HUM; 66.8% for AEM; 57.2% for OFM). In addition to an increase in explained deviance, the residual autocorrelation (weak but observed in the HUM and OFM) was no longer apparent in the AEM (ACF plots; Figure 2.5, Figure 2.6,Figure 2.7), in which SST was substituted by geographic position (although the auto-correlation in the residuals of the HUM and OFM were not high and required no further action; see Wood [2006] for concerns about residual autocorrelation of GAMs). The bivariate smooth for easting/northing included in the AEM is likely acting as a proxy for unmodelled environmental, behavioural or individual characteristics. For example, because it was highly correlated with SST, which was not included in the AEM, easting/northing may be representing not only SST but also some other environmental feature(s). This may explain the increase in percentage of explained deviance when SST is substituted by easting/northing in the AEM.

Shelter (a combination of SST and wind speed) was created as an environmental feature that could be important to whales that are calving, for example, to represent conditions that may be related to energy saving for the calf (Corkeron and Connor, 1999). Because the effects of wind speed on detectability have been accounted for in the estimation of the detection probability, no confounding with the effects of wind in the shelter variable is expected. The response variables in the detection function model and the habitat use/abundance estimation spatial models are completely different. In the detection process, it is the perpendicular distance (in relation to the trackline); in the spatial models the response variable is the estimated abundance. Furthermore, wind speed may influence both the detectability of animals and how animals use their habitat, which is supported by the present results. Indeed, a major advantage of DSMs using data from distance sampling surveys is that the effects of variables on detectability and on abundance can be teased apart.

The DSM approach permitted inference and extrapolation from the AEM to the area not surveyed in 2012 (Bortolotto et al., 2016a), resulting in a 2012 abundance estimate for a larger part of the breeding ground distribution than would otherwise be available. The lack of data to the north of Salvador in 2012 implies that the effect of the bivariate smooth

for easting/northing on the predictions for that area is largely influenced by data from 2008. However, the other variables retained in the model were responsible for the large majority of the explained deviance, as illustrated by the percentage of explained deviance of the HUM (54.1%), so this is not considered to be an important limitation for inferences about abundance here.

Model-based abundances for humpback whales breeding off the coast of Brazil (14,264, CV = 0.084 for 2008; 20,389, CV = 0.071 for 2012) were estimated to be close to those computed by design-based methods (16,410, CV = 0.228 for 2008; 19,429, CV = 0.101 for 2012; Bortolotto et al., 2016a). This similarity could be expected because both estimates were derived from the same data, although uniform coverage probability, a requirement for adequate design-based estimates (Buckland et al., 2015), was not attained (Figure 2.1). The higher precision in the present model-based abundance estimates (CV = 0.084 vs. 0.228 for 2008; CV = 0.071 vs. 0.101 for 2012) is mainly because the covariates explained some of the variability in the data, demonstrating the value of the analysis.

2.4.2 Habitat use

The main reasons for SST to be considered an important factor in explaining the distribution of migratory whales in their breeding grounds are likely related to presence of calves, which are not as efficient in conserving their body temperature as older animals (Corkeron and Connor, 1999). SST was the most important variable selected in the HUM, and it was highly correlated with geographic position (northing/easting). The overall relationship between whale density and SST was positive, peaking at 24 to 25°C. This result for SST may reflect habitat selection of calving females for the reason stated above. The habitat use of North Atlantic right whales in their calving grounds off the south-eastern US was also observed to be strongly related to SST (Keller et al., 2006); however, differences in species characteristics (e.g., latitudinal range) should be taken into account in any comparison. Trudelle et al. (2016) did not find a relationship between SST and humpback whale movements in their Madagascar coastal breeding area, possibly because of the relatively low variation in SST in the area. Although a temporal change in distribution was not supported by the models, long-term monitoring should provide

important insights to this, because the effects of climate change (Walther et al., 2002), for example, may impact the distribution of marine animals.

Shelter, which incorporated SST, was consistently retained in the spatial models and therefore can be considered an important factor in explaining this population's distribution in the breeding area. The fitted relationship for this covariate suggests that relatively light and moderate surface winds had a significant positive effect on density, when the water was relatively colder. Because wind speed was not selected in the spatial models, results suggest that wind may be an important habitat feature for WSA humpback whales only when the water temperature is relatively cool. A possible reason is that because temperature is one of the most important features for these animals in the area, they tolerate a range of wind speeds beyond their preferred wind speed range when SST is relatively warmer. As mentioned above, because calves may benefit from an environment where they can save body energy reserves, calm conditions at the water surface are likely preferable for calves to swim and to surface to breathe (Taber and Thomas, 1982; Cartwright et al., 2012). In a daily-scale study of habitat use, Félix and Botero-Acosta (2011) found that mother-calf humpback whale pairs in Ecuador preferred shallower waters during the afternoon hours, when wind speeds in the area tended to increase and the sea tended to become rougher. The combination of water temperature and wind at the surface seems to be an important factor for WSA humpback whale habitat selection in breeding grounds. That whale groups presenting calves may prefer sheltered waters has been suggested before (Craig et al., 2014). However, Rayment et al. (2015) is probably the only previous study that incorporated a variable to explicitly represent shelter in habitat use models for breeding migratory whales. These authors investigated the influence of shelter in the breeding distribution of right whales and found that wave exposure and distance to shelter (defined as areas with lower wind exposure) influenced habitat selection of right whale groups with calves. Given the present results for shelter and the strong relationship estimated between whale density and sea water temperature, climate change may have severe consequences for humpback whales off Brazil. If, as a consequence, sea water temperature increases and storms become more frequent (Webster et al., 2005), animals will be facing a more challenging environment in the future.

It is still unclear which environmental features really represent shelter for breeding whales and how this may vary among different species. Martins et al. (2001) showed that the occurrence of WSA humpback whale groups containing calves increased with the proximity to the Abrolhos Archipelago, which may represent shelter for these animals, with the presence of the archipelago perhaps creating a calmer environment. Also, Zerbini et al. (2004) observed that WSA mother-calf groups were more frequently found closer to the shore than other group types off the north-eastern coast of Brazil. Present results add to this discussion of which environmental variables may combine to create a sheltered environment that benefits migratory whale species in their breeding grounds. While several other covariates could have been included or combined to create a spatial covariate to represent shelter (e.g., speed and direction of ocean currents), the simple combination that was presented here for shelter permitted easy interpretation of model results. A complicated combination of several covariates would likely produce results that would be difficult to interpret biologically.

Relationships between whale density and environmental covariates revealed by the present models are consistent with what could be expected for mothers, which may show a preference for secure environments for the development of their calves in sheltered waters. However, as noted by Trudelle et al. (2016), while the movements of female humpback whales in a breeding area off the coast of Madagascar are influenced by environmental features such as depth and distance to the shore, male movements are probably more influenced by social factors, such as female occurrence. Despite the fact that their distribution may also be influenced by the presence of other males (Herman, 2017), adult male humpback whales are indeed likely to seek receptive females, not those that are about to or have just given birth. Craig et al. (2014) suggested that calving female humpback whales in Hawaii prefer shallow waters where chances of being harassed by males are supposedly lower; their habitat selection is likely driven primarily by avoidance of males, which may prefer deeper waters for breeding interactions (Smultea, 1994). Humpback whale groups containing calves have been found significantly more frequently in shallower waters than groups without calves in Brazilian breeding grounds (Martins et al., 2001; Zerbini et al., 2004). Thus, bathymetric features may also be related to what may represent shelter for these animals.

Overall, this discussion highlights the importance of having data on the sex and reproductive status of individuals and not only on environmental features to understand the distribution of large whales in breeding areas. For example, bathymetry was not considered as part of shelter to facilitate interpretation of results, but if individual data were available it could be informative to investigate a wider range of covariate combinations representing shelter in models of habitat use. Future studies could also investigate in detail the conditions of the marine environment in areas surrounding the Abrolhos Archipelago, because the presence of coral reefs may be related to (or contribute to) shelter from rough water (Lindsay et al., 2016).

The positive relationship between whale density and distance to both the coast and the continental shelf break could mean that humpback whales off the coast of Brazil prefer to be in the middle part of the shelf, or that they avoid the shelf boundaries. Trudelle et al. (2016) suggested that the distance to coast was one of the most important factors affecting the movement patterns of female humpback whales off the Madagascar breeding grounds, and other studies have shown that calving humpback whales are associated with areas close to the shore (Martins et al., 2001; Zerbini et al., 2004; Félix and Botero-Acosta 2011). Avoidance of the shelf edge could be in response to the risk of predation by large predators in offshore waters, such as large shark species (Smultea, 1994). Areas too close to the shore could be avoided because they are too shallow for swimming (Oviedo and Solís, 2008) or because of disturbances that were not considered here, such as noise from human activities.

The estimated negative effect on predicted whale numbers of current speeds greater than 0.2 m s^{-1} is not very well supported by the data (95% confidence interval widens with increasing current speed). The peak of whale density in the estimated relationship with current speeds indicates that whales in the area may preferentially select habitats where current speeds are around 0.2 s m^{-1}. Results presented in Trudelle et al. (2016) support the importance of the current for large whales in breeding areas, since differences in current speed between shelf and oceanic waters influenced the movement patterns of humpback whales in their breeding area off Madagascar. Whales of both sexes swam faster in slower currents, and the authors of that study suggested that when animals are engaged in mate-searching-related movements close to the coast, the current speed

probably does not have an important effect. Therefore, data on the behavioural status and/or movements of individual animals are likely needed to better understand the effects of current speed on habitat use of humpback whales off the coast of Brazil. In addition, the resolution of this covariate (5 d and 0.33 × 0.33° of latitude/longitude; Table 2.1) was likely unable to capture fine-scale variability, particularly around complex coastlines.

2.4.3 Oil and gas production fields

Because this increasing population was severely depleted before the intense development of coastal activities, such as those related to hydrocarbons extraction, fishing and boat traffic, it is expected that conflicts will arise from whales re-occupying areas that they previously inhabited (Andriolo et al., 2010). As presented here, habitat use studies can help understanding of how whales are distributed in relation to their surrounding environment. It is fundamental that human-related stressors are considered when studying habitats constantly changed by human activities.

In this chapter, the proximity to oil production fields was considered as a candidate variable to explain variation in humpback whale density in coastal Brazilian waters. However, the direct effect of the presence of oil installations on the distribution of whales, or a behavioural response to oil exploration-related activities, could not be tested. What is presented here is simply evidence for a relationship, which indicates that the distributions of both humpback whales and oil production fields in the area follow similar patterns. To investigate animal response, different approaches and data would be required. It is not possible to conclude that the current distribution patterns are a response to the presence of oil and gas installations, because of multiple reasons. Firstly, no information on animals' distribution is available from before oil exploration activities started. Secondly, estimated relations between whale density and covariates do not necessarily indicate causation, as happens in every regression modelling (Amr and Stamboliyska, 2016). Finally, because very little has changed in the distribution of oil fields within the period considered here (Figure 2.2), an interaction term between year and geographic position was not investigated here. Before potential redistribution of animals can be related to changes in man-made structures in the area, more evident changes in structures configuration and/or number must happen. A change in oil fields

over time and independent of changes in other variables would present an opportunity to investigate avoidance by whales in a quasi-experimental framework.

The placement of oil production fields is not random, but follows a process that considers several factors, including the potential impact to the natural environment. For example, in 2003 a study on the potential impact of oil operations on the Abrolhos Bank (Marchioro et al., 2005), resulted in the cancelation of 243 oil exploration blocks in the Bank, which were offered at the 5th bidding round by ANP (ANP, 2017). That decision was supported by knowledge of the high densities of whales in the area (Marchioro et al., 2005). Given that the configuration of oil fields present in the survey area was partially influenced by the presence of animals, the distance to these installations could be expected to be positively related to whale concentrations. In other words, oil fields could be expected to be far from where whales are more frequently found which is supported by the OGM results.

Another explanation for the negative relation between the distance to oil fields and whale density for distances greater than 100 km may be the fact that the survey area forms roughly a long strip, which is oriented latitudinally (Figure 2.1). Because a higher concentration of animals was found in a relatively confined region within the survey area (south of the Abrolhos Archipelago; Figure 2.11), this relation may be simply reflecting the latitudinal distribution of the fields. This also explains the peak abundance at around 100 km (Figure 2.10), which is the distance from the higher concentration area to the nearest oil field, close to the coastline of Espírito Santo State (Figure 2.2). The effects of the physical presence of oil production fields should be very minimal 100 km away, so there is likely other covariate related to space being represented here, such as latitude and longitude.

Here, the production fields are taken to be the delimited areas where the oil activity was in the production phase (ANP, 2017) during the time that the surveys were conducted. Production fields in the marine portion of the sedimentary basins covered in the surveys were considered here. Thus, some fields beyond the survey limits were included (Figure 2.2). However, there are several other potential sources of impact from oil operations, such as seismic surveying, ship traffic and the risk of oil spills (Chapter 4; Engel et al.,

2004; Martins et al., 2013; Rossi-Santos, 2015). What is also not considered here is the clustering of oil fields. For example, Campos basin (around Rio de Janeiro State), in the south of the study area (Figure 2.2), has a relatively higher density of fields than other areas. All that considered, it is not possible to conclude what the presence of oil fields represents for humpback whales, apart from a potential source of disturbance. However, the high concentration of oil fields in the south of the area (Figure 2.2) may present an important disturbance for animals initiating migration (Zerbini et al., 2006b). To further investigate the potential impact that oil activities may have on this population, areas of higher risk of impact from oil spills were identified, as described in Chapter 4.

2.4.4 Implications for conservation and management

The predicted distributions support previous work showing that WSA humpback whales are strongly related to features in the Abrolhos Bank region during their breeding season in coastal waters of Brazil (Siciliano, 1997; Andriolo et al., 2010; Wedekin, 2011; Martins et al., 2013; Pavanato et al., 2017). However, other areas also had relatively high predicted densities, such as near Salvador and off the coasts of Sergipe and Alagoas States (Figure 2.13 and Figure 2.14). Little is known about their distribution or habitat use in these areas (Zerbini et al., 2004; Baracho-Neto et al., 2012), but relatively recent observations indicate that the distribution of WSA humpback whales in Brazil may be broader than previously recognized (e.g., Wedekin et al., 2014; Bortolotto et al., 2016c; Pavanato et al., 2017).

The Abrolhos Archipelago is included in the Abrolhos Marine National Park (Figure 2.15), which is a national 'Conservation Unit' area of 880 km^2 (ICMBio, 2017). According to the Brazilian Ministry of Environment (www.mma.gov.br) this is a federal conservation unit of 'integral protection' where only scientific research and educational, recreational and small-scale ecotourism activities are permitted. All of these activities are regulated by the Chico Mendes Institute for Biodiversity Conservation (ICMBio), the federal body responsible for protected areas in Brazil. Commercial activities are therefore mostly limited to those related to small-scale ecotourism. The nearby Environmental Protection Area of Ponta da Baleia is regulated by Bahia State and is in the category of 'sustainable use area' (INEMA, 2017). These protected areas cover a very small portion of the area predicted to have the highest concentration of humpback whales (Figure 2.15).

Present results support the conclusions of Castro et al. (2014), who used satellite-tracked movement data to show that MPAs only cover a very small portion of the areas most used by WSA humpback whales in their breeding grounds, representing only 0.64% of the population's wintering habitat (Castro et al., 2014).

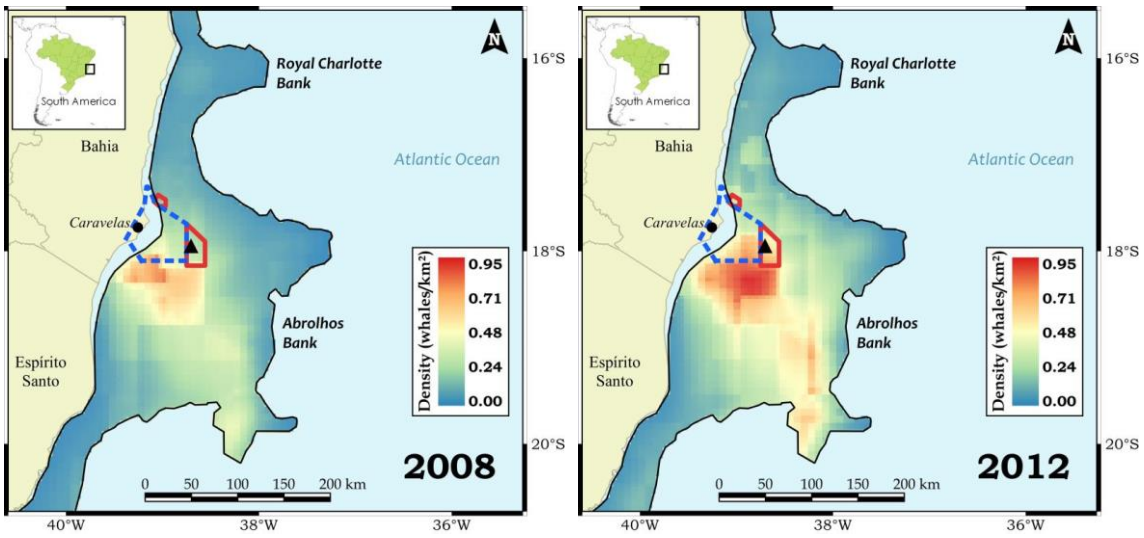


Figure 2.15. Density surface maps for 2008 and 2012 for the Abrolhos Bank region. Predictions were made with the abundance estimation model (AEM). Black triangles indicate the location of the Abrolhos Archipelago. Red polygons represent the Abrolhos Marine National Park, and dashed blue polygons represent the Ponta da Baleia marine protected area.

The Abrolhos Bank is a region of high biodiversity (Werner et al., 2000), and expanding the area under protection, restricting commercial activities, could benefit not only cetaceans but also other marine organisms, such as the unique coral reefs in the area (Francini-Filho and Moura, 2008). Because most humpback whale births are expected to occur on or near Abrolhos Bank (Martins et al., 2001), expanding the protected area during the period when whales are consistently present (winter−spring), could reduce the risk of anthropogenic impact, especially for calves that are more vulnerable to disturbance (Schaffar et al., 2013). To conserve marine species in the area, past management actions have included the cancellation of seismic and other oil and gas exploitation activities on the Bank during the humpback whale breeding season (Engel et al., 2004; Marchioro et al., 2005). However, there is increasing interest from the oil and gas industry to explore for oil on the Bank (http://app.anp.gov.br). Group composition was not included in this study but because young animals are more vulnerable to stressors (Schaffar et al., 2013;

Ott et al., 2016; Dunlop et al., 2017), future studies aiming to provide information for conservation should investigate the distribution of different group types at a finer scale and include potential stressors and displacement factors associated with human presence in the marine environment, with special attention to the Abrolhos Bank region.

Abundance estimates presented here (14,264, CV = 0.084 for 2008 and 20,389, CV = 0.071 for 2012) provide additional confirmation that the WSA humpback whale population is growing (Zerbini et al., 2011). A new population status assessment in the framework of Zerbini et al. (2011) is investigated in Chapter 5, which considers the present results and new information on population growth rate (Wedekin et al., 2017) to provide an updated understanding of this population's recovery, more than four decades after whaling ceased in 1973 in this area.

It is important that efforts to monitor potential threats are intensified because current knowledge about the impacts of human activities is very limited (Bezamat et al., 2015; Bortolotto et al., 2016c; Ott et al., 2016). To evaluate adequately the need for improvement or adjustment of current conservation strategies and management actions, such as enhancing protection in the area (Castro et al., 2014), it is essential to assess the conservation status of WSA humpback whales and to assess current and future potential impacts on the population. The distribution results presented here may also be used in evaluating areas of higher risk for this population by investigating sources of impact by human-related activities in the areas predicted to be most used by the animals. This was partially investigated in Chapter 4, where a simple method for risk assessment from oil spill impacts on these whales is described.

CHAPTER 3

DISTRIBUTION MODELS FROM HUMPBACK WHALE SATELLITE TELEMETRY DATA

ABSTRACT

Investigating habitat use of migratory whale species in breeding areas is challenging and different survey methods may be complementary to its understanding. For the humpback whale population found in Brazil during winter and spring, there is information on habitat use from survey data. However, similar information from tracking data is currently missing, which may be informative. In this chapter, the distribution of humpback whales in breeding grounds of Brazil was modelled using individual tracking data, from satellite telemetry, and inferences were compared with those from line transect data modelling to verify differences and/or similarities in both methods. Distribution models were fitted to locations along interpolated tracks using the Generalized Estimation Equation framework, modelling presence and pseudo-absence as smooth functions of continuous covariates with flexible spline knot location. Covariates considered were latitude and longitude, SST, current and wind speeds close to the surface, distances to shelf break and to coast, depth and sea bottom slope. A factor variable representing shelter (a combination of SST and wind speed values) was also considered. Two modelling exercises were conducted: a habitat use model (HUM) that included all covariates in the candidate set except latitude and longitude; and a distribution model (DIM) that also included latitude and longitude. The selection of candidate covariates was based on the model *QICu* score and the statistical significance of covariates. Covariates included in the final HUM were SST, distance to the coast and to the shelf break, current and wind speeds, and shelter. In the DIM, latitude and longitude replaced SST because they were correlated; other covariates retained in the final model were current speed, and distances to both shelf break and coast. The relationships between density and environmental covariates from models of tracking data corroborate those from models of line transect data. Predicted maps were also similar, indicating high concentration around the Abrolhos archipelago and to the south. Results indicate that habitat use characteristics of this population in the area can be

inferred similarly from models of different data types. For spatial ecology and habitat use studies in large scales, spatial coverage can be improved by modelling data from line transect and telemetry tracking if available for complementary areas.

3.1 INTRODUCTION

Habitat use and distribution are the outcome of a combination of animals exploiting available resources to meet their life requirements, their physiological restrictions and the effects of the environment (Aarts and Matthiopoulos, 2010), including effects from human activities, which may be a source of disturbance. Management of human activities to avoid impact on animal populations can be better informed through knowledge of where animals concentrate and how they use the environment. In this context, distribution models are useful tools to support the identification of areas that require management, such as protected areas (Runge et al., 2015).

Different methods exist to estimate distribution and investigate habitat use for animal populations, and the outputs strongly depend on the nature of the information collected (Redfern et al., 2006; Aarts et al., 2008). Line transect sampling is designed to estimate abundance and the data are commonly used to infer habitat use through the application of spatial models (e.g., density surface models, DSMs; Miller et al., 2013). Such surveys are conducted by sampling lines distributed in a survey area, recording detected animals or groups as well as their distances from the transect to allow effective search area to be estimated (Buckland et al., 2001). The sampling unit is therefore space (i.e., the area searched around the transect) within which animals are either present or absent and may potentially be observed.

Another method to investigate how marine mammals use their habitat is through tagging individual animals with animal-borne tracking devices, which has been facilitated by the development of equipment and analysis tools (Aarts et al., 2008; Russell et al., 2016; Jonsen, 2016; Trudelle et al., 2016). In contrast to line transect surveys, the sampling unit in telemetry tagging is the individual animal and the data are the locations visited by the tagged animals along individual tracks. Therefore, the space that is sampled is only that which is visited by tagged animals, meaning that data from places not visited by them are unavailable. However, there are several issues to consider when using telemetry data for

that purpose (Aarts et al., 2008). For example, the distribution of locations can be much influenced by the distribution of locations of tag deployment, creating a tagging location bias (Block et al., 2011). Another important aspect of telemetry tracking data is the autocorrelation in locations along tracks, since naturally this type of data is collected in sequence in time and space (Aarts and Matthiopoulos, 2010). Moreover, usually the locations are provided from tracking of a small portion of the population for which inferences are intended, meaning that although sample size for locations are usually large, sample sizes for individuals are relatively small.

Presence-absence approaches can be used to infer distribution of marine mammals from tracking data however, because of the lack of information on real absences (e.g., Pirotta et al., 2011), "pseudo-absences" (arbitrarily generated points, usually randomly placed across the survey area) must be generated to represent the available background habitat (Aarts et al., 2008; Russell et al., 2016; Jones et al., 2017). However, despite the method being commonly used in distribution studies, there are several analytical issues to consider. When using pseudo-absences, the number of absences generated to include in the models is controlled by the user; there is some disagreement among researchers about the pseudo-absence/presence ratio (Manly et al., 2002; Keating and Cherry, 2004; Aarts et al., 2008) that ensures that the information about the available, but not visited, habitat is reliable. Pseudo-absences can potentially be placed where animals were actually present, which is suggested to not be a very important problem for large scale studies (Aarts et al., 2008). A decision about how to define the area of inference is required: some studies used a percentage of kernel density estimates to represent habitats frequently used by the animals (e.g., Pendoley et al., 2014; Mei et al., 2017; Thorne et al., 2017); this percentage is, again, user-controlled. Because spatial data obtained through monitoring moving animals are naturally correlated in space and time, due to presences being sequences of locations in a time-series, such correlation must be dealt with or taken into account for habitat use inferences to be reliable (Aarts et al., 2008).

Alternative approaches to investigate habitat use from telemetry/tagging data include inferring animal behaviour, which can be related to specific biological and environmental features (McClintock et al., 2015; Jonsen, 2016). There are also methods to study spatial distribution from tagging data which do not require creating pseudo-absences, such as

point process models (Johnson et al., 2013) or by considering animal tracks in a grid over the study area as a result of a Markov process (Whitehead and Jonsen, 2013).

Most humpback whale populations are recovering after the end of global commercial whaling (Zerbini et al., 2010) and the population that mates and calves in coastal waters of Brazil during winter and spring is no exception (Bortolotto et al., 2016a; Wedekin et al., 2017). However, the coincident increase in human activities in Brazilian coastal waters, especially those related to oil and gas production (Chapter 2; ANP, 2017; Bortolotto et al., 2017), means that there is a need to understand how these activities may affect distribution and habitat use to inform whether management actions may be necessary to avoid impact on this population. Distribution of this population in the area was investigated in the past using descriptive analysis (Zerbini et al., 2006; Martins et al., 2001; 2013; Castro et al., 2014; Gonçalves et al., 2018), from difference in densities within line transect survey blocks (Andriolo et al., 2010) or for very restricted areas (Martins et al., 2001; Gonçalves et al., 2018). More recently, analysis of line transect data to investigate the distribution of humpback whales off the coast of Brazil have shown that density was strongly related to temperature and bathymetric features (Chapter 2; Bortolotto et al., 2017; Pavanato et al., 2018). Telemetry data for the species in the area was used to investigate relative usage of protected areas (Castro et al., 2014) and movements (Zerbini et al., 2006).

Since line transect surveys rely on sampling the area and telemetry tagging on sampling individual animals, the methods provide information of different kinds. The objective in this chapter was to compare outputs from spatial distribution models, applied to telemetry data, with those from Chapter 2, in which line transect data were used. Because different types of data for studying species' spatial ecology may be available, it is important to explore whether similar inferences about distribution and habitat use should be expected. For example, data obtained through different methods may be available from different regions within an area of interest, so the spatial extent of information may be improved. Thus, management actions may be better informed about distribution patterns and habitat use of populations when data from different methods can be used.

3.2 MATERIALS AND METHODS

The area of inference in this study was defined by adapting the survey area previously used in Chapter 2, considering the extent of available environmental variables and tagging locations. In presence/pseudo-absences studies the area of inference, or survey area, is considered as the habitat available for the animals (Aarts et al., 2008). Thus, because tagging activities were not distributed evenly across the study area (Figure 3.1), the northernmost portion of the original area was excluded. The resulting survey area (Figure 3.1) was assumed to be the area available for the studied animals.

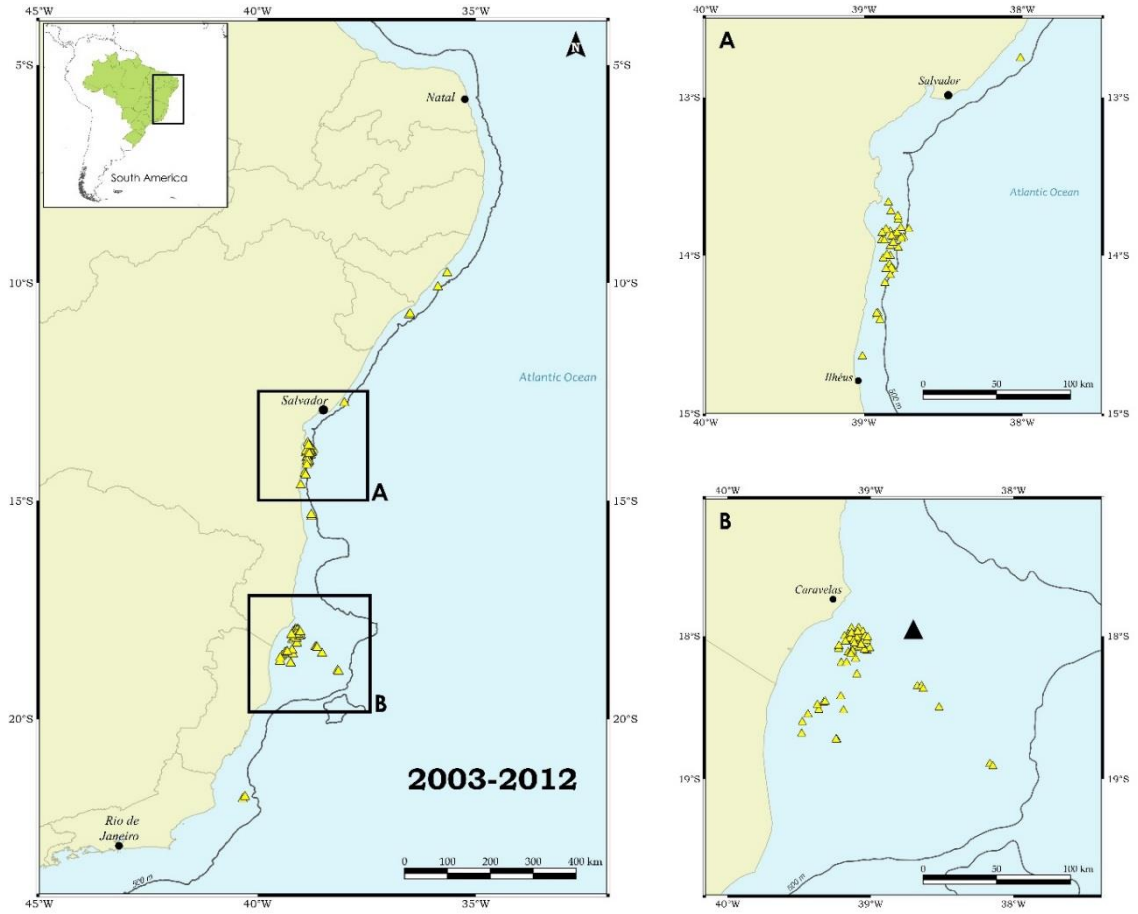


Figure 3.1. Tagging locations of humpback whales along the Brazilian coast. Animals were tagged from 2003 to 2012. A: Camamú region, Bahia state. B: the Abrolhos Archipelago region. The archipelago location is indicated by a black triangle.

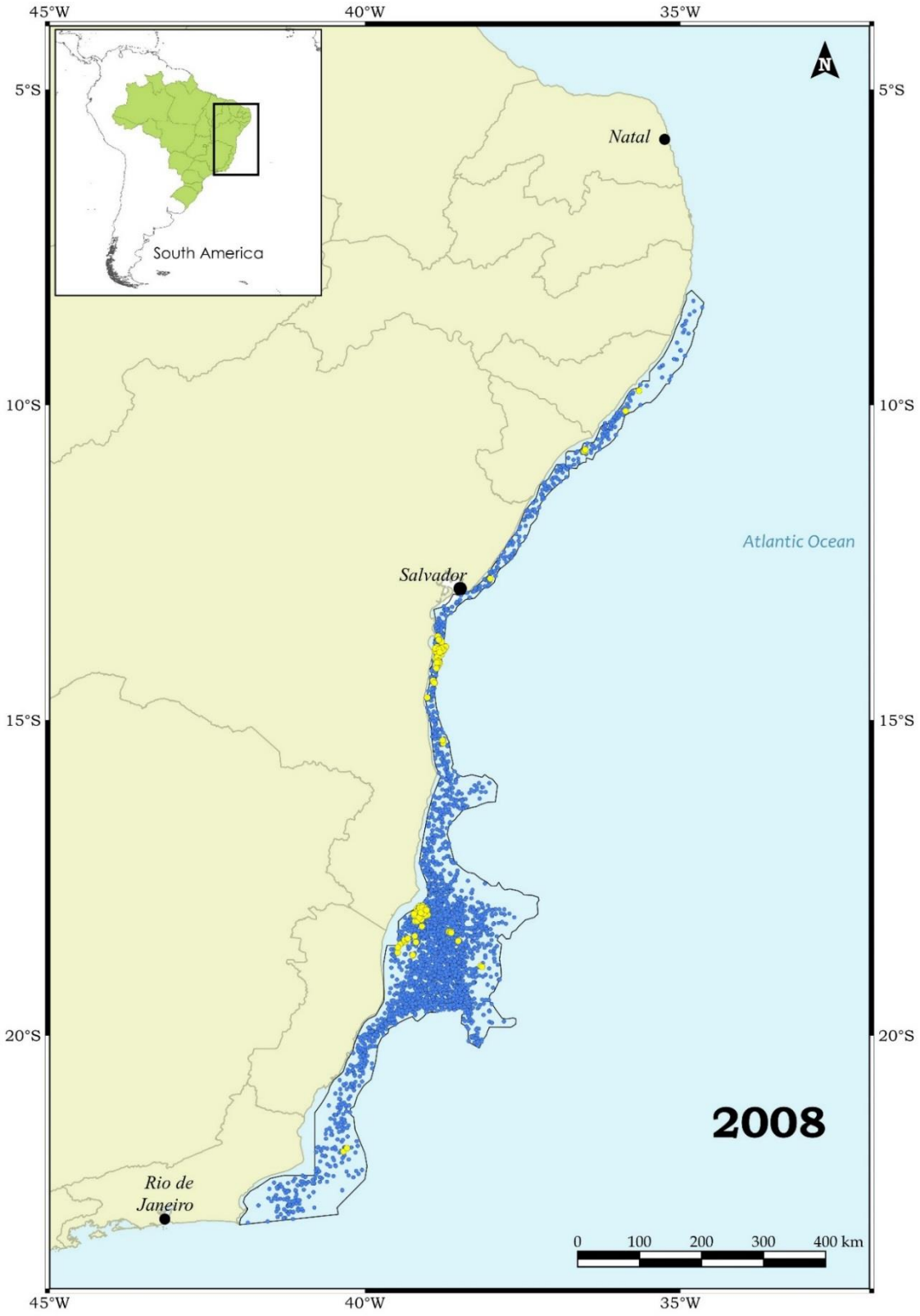


Figure 3.2. Locations used in the distribution and habitat use analysis (blue) and tagging locations (yellow). Tagging locations (Figure 3.1) are presented again here to illustrate the balance between the tagging and presences across the area of inference (black line).

3.2.1 Data acquisition: whale tracking

From 2003 to 2012, satellite-linked telemetry tags were attached to adult humpback whales along the coast of Brazil. Tagging operations occurred every year between August and December except for 2004, when no tagging happened (Table 3.1). Tagging operations were conducted in two ways, depending on where the research team was based during activities. From 2003 to 2007 and from 2009 to 2012, the research team was based on land, working at sea every day from power-boats. In 2008 and 2012, the research team was based on the R/V Atlântico Sul (*Universidade Federal do Rio Grande*, FURG), during research cruises that surveyed the Brazilian continental shelf for multiple research purposes (Chapter 2; Bortolotto et al., 2016a). Research cruises collected line transect data for abundance and distribution investigations, and also tagged animals in a wider area than otherwise would have been possible from land-based operations. In 2012, ship-based tagging operations were not completed as planned because of adverse weather conditions and were continued later that year from a land-based station.

Table 3.1. Information on whale tagging. (Whale ID = unique identification of tag; Area = area of tagging in the coast of Brazil; Group = group composition; Role = social role of whale in the group; Cycle = programmed cycle for attempting transmissions; Transm. = programmed number of attempted transmissions within "Interval" and "Cycle"; Interval = interval within a day, i.e. 24 hours, when transmissions were programmed to be attempted; eod = every other day; all = every day; e4d = every four days; U = undetermined).

Whale ID	Date	Latitude	Longitude	Area	Group	Role	Sex	Transm.	Cycle	Interval
21810.03	18/10/2003	-18.683	-39.483	Abrolhos	mother, calf, escort	escort	M	300	eod	6-21
27261.03	18/10/2003	-18.724	-39.237	Abrolhos	mother, calf, escort	escort	M	300	eod	6-21
24640.03	19/10/2003	-18.548	-39.438	Abrolhos	single	adult	F	300	eod	6-21
27259.03	19/10/2003	-18.516	-39.362	Abrolhos	mother, calf, three adults	escort	M	300	all	6-21
20162.03	19/10/2003	-18.481	-39.372	Abrolhos	mother, calf, three adults	escort	M	300	eod	6-21
24642.03	27/10/2003	-18.520	-39.190	Abrolhos	mother, calf	mother	F	300	eod	6-21
21809.03	27/10/2003	-18.422	-39.209	Abrolhos	mother, calf	mother	F	500	all	0-23
21792.03	27/10/2003	-18.464	-39.329	Abrolhos	mother, calf, escort	escort	M	500	all	0-23
20687.03	27/10/2003	-18.459	-39.319	Abrolhos	mother, calf, escort	mother	F	500	all	0-23
21800.03	28/10/2003	-18.724	-39.242	Abrolhos	mother, calf	mother	F	500	all	0-23
21791.03	28/10/2003	-18.602	-39.479	Abrolhos	mother, calf	mother	F	500	all	0-23
26712.05	11/10/2005	-18.012	-39.068	Abrolhos	mother, calf, escort	mother	F	300	Out-Nov: all; Dec-: e4d	7-22
7617.05	11/10/2005	-18.021	-39.050	Abrolhos	mother, calf, escort	mother	F	300	Out-Nov: eod; Dec-: e4d	0-23
7618.05	11/10/2005	-18.018	-39.065	Abrolhos	mother, calf	mother	F	300	Out-Nov: eod; Dec-: e4d	0-23
27261.05	12/10/2005	-17.946	-39.134	Abrolhos	mother, calf	mother	F	300	Out-Nov: eod; Dec-: e4d	7-22
27259.05	16/10/2005	-18.062	-39.111	Abrolhos	mother, calf	mother	F	300	Out-Nov: eod; Dec-: e4d	7-22
37229.05	16/10/2005	-18.083	-39.110	Abrolhos	mother, calf, escort	mother	F	300	Out-Nov: eod; Dec-: e4d	7-22
37231.05	16/10/2005	-18.068	-39.107	Abrolhos	mother, calf, escort	mother	F	300	Out-Nov: eod; Dec-: e4d	7-22
37234.05	16/10/2005	-18.042	-39.099	Abrolhos	mother, calf, escort	mother	F	300	Out-Nov: eod; Dec-: e4d	7-22
24641.05	19/10/2005	-18.103	-39.026	Abrolhos	mother, calf	mother	F	300	Out-Nov: all; Dec-: e4d	7-22
27258.05	19/10/2005	-18.037	-39.151	Abrolhos	mother, calf	mother	F	300	Out-Nov: eod; Dec-: e4d	7-22
10946.05	19/10/2005	-18.009	-39.156	Abrolhos	mother, calf	mother	F	300	Out-Nov: eod; Dec-: e4d	0-23

33000.05	19/10/2005	-18.078	-39.083	Abrolhos	mother, calf	mother	F	300	Out-Nov: eod; Dec-: e4d	0-23
33001.05	19/10/2005	-18.097	-39.040	Abrolhos	mother, calf	mother	F	300	Out-Nov: eod; Dec-: e4d	0-23
37229.06	12/10/2006	-17.988	-39.156	Abrolhos	two adults	adult	U	300	all	7-22
33000.06	12/10/2006	-18.001	-39.182	Abrolhos	mother, calf, escort	mother	F	300	all	7-22
27259.06	13/10/2006	-17.996	-39.129	Abrolhos	mother, calf	mother	F	300	all	7-22
37282.06	18/10/2006	-18.035	-39.154	Abrolhos	four adults	adult	U	300	eod	7-22
37288.06	20/10/2006	-18.017	-39.159	Abrolhos	mother, calf	mother	F	300	eod	7-22
50682.06	25/10/2006	-18.004	-39.155	Abrolhos	mother, calf	mother	F	300	eod	7-22
33001.06	25/10/2006	-17.993	-39.087	Abrolhos	mother, calf	mother	F	300	all	7-22
37236.06	26/10/2006	-18.187	-39.206	Abrolhos	mother, calf	mother	F	300	eod	7-22
37230.06	26/10/2006	-18.185	-39.169	Abrolhos	mother, calf	mother	F	300	all	7-22
42521.06	26/10/2006	-18.268	-39.096	Abrolhos	U	U	U	300	eod	7-22
37234.06	28/10/2006	-18.022	-39.131	Abrolhos	mother, calf, escort	mother	F	300	eod	7-22
42521.07	08/09/2007	-18.011	-39.099	Abrolhos	mother, calf	mother	F	300	all	0-23
60007.07	08/09/2007	-18.012	-39.096	Abrolhos	mother, calf	mother	F	300	all	0-23
27261.07	12/09/2007	-17.979	-39.083	Abrolhos	mother, calf, escort	mother	F	300	all	0-23
37231.07	12/09/2007	-17.984	-39.093	Abrolhos	mother, calf, escort	escort	M	300	all	0-23
37286.07	14/09/2007	-17.998	-39.140	Abrolhos	mother, calf	mother	F	300	all	0-23
37288.07	14/09/2007	-18.005	-39.147	Abrolhos	mother, calf, escort	mother	F	300	all	0-23
50682.07	15/09/2007	-18.038	-39.176	Abrolhos	mother, calf, escort	mother	F	300	all	0-23
50686.07	15/09/2007	-18.117	-39.140	Abrolhos	mother, calf, escort	mother	F	300	all	0-23
42525.07	16/09/2007	-18.113	-39.157	Abrolhos	mother, calf	mother	F	300	all	0-23
50687.07	16/09/2007	-18.158	-39.106	Abrolhos	mother, calf, escort	mother	F	300	all	0-23
60004.07	16/09/2007	-18.074	-39.225	Abrolhos	single	adult	M	300	all	0-23
87759.08	27/08/2008	-18.499	-38.521	Abrolhos	two adults	adult	M	300	Sep-Nov: all; Dec-: eod	7-22
87760.08	28/08/2008	-15.362	-38.761	Bahia	two adults	adult	U	300	Sep-Nov: all; Dec-: eod	7-22
87761.08	28/08/2008	-15.316	-38.769	Bahia	two adults	adult	M	300	Sep-Nov: all; Dec-: eod	7-22
87762.08	29/08/2008	-12.752	-38.012	Bahia	two adults	adult	M	300	Sep-Nov: all; Dec-: eod	7-22

87763.08	30/08/2008	-9.772	-35.654	Alagoas	mother, calf	mother	F	300	Sep-Nov: all; Dec-: eod	7-22
87764.08	06/09/2008	-10.094	-35.867	Alagoas	mother, calf	mother	F	300	Sep-Nov: all; Dec-: eod	7-22
87765.08	07/09/2008	-10.737	-36.513	Sergipe	mother, calf, four adults	adult	F	300	Sep-Nov: all; Dec-: eod	7-22
87767.08	07/09/2008	-10.708	-36.506	Sergipe	mother, calf	mother	F	300	Sep-Nov: all; Dec-: eod	7-22
87766.08	08/09/2008	-10.732	-36.527	Sergipe	mother, calf, four adults	adult	M	300	Sep-Nov: all; Dec-: eod	7-22
87768.08	12/09/2008	-14.381	-38.918	Bahia	mother, calf, escort	mother	F	300	Sep-Nov: all; Dec-: eod	7-22
87769.08	12/09/2008	-14.366	-38.912	Bahia	two adults	adult	U	300	Sep-Nov: all; Dec-: eod	7-22
87770.08	12/09/2008	-14.368	-38.920	Bahia	mother, calf, escort	escort	M	300	Sep-Nov: all; Dec-: eod	7-22
87773.08	12/09/2008	-14.639	-39.010	Bahia	mother, calf	mother	F	300	Sep-Nov: all; Dec-: eod	7-22
87775.08	16/09/2008	-18.351	-38.673	Abrolhos	four adults	adult	F	300	Sep-Nov: all; Dec-: eod	7-22
87776.08	16/09/2008	-18.350	-38.644	Abrolhos	two adults	adult	F	300	Sep-Nov: all; Dec-: eod	7-22
87777.08	16/09/2008	-18.369	-38.632	Abrolhos	mother, calf, three adults	mother	F	300	Sep-Nov: all; Dec-: eod	7-22
87774.08	17/09/2008	-18.894	-38.166	Abrolhos	two adults	adult	U	300	Sep-Nov: all; Dec-: eod	7-22
87778.08	17/09/2008	-18.910	-38.145	Abrolhos	two adults	adult	M	300	Sep-Nov: all; Dec-: eod	7-22
87771.09	14/09/2009	-13.901	-38.887	Bahia	five adults	adult	U	300	Aug-Oct: all, Nov-: eod	5-10, 17-22
87773.09	15/09/2009	-13.773	-38.784	Bahia	single	adult	U	300	Aug-Oct: all, Nov-: eod	5-10, 17-22
87774.09	15/09/2009	-14.125	-38.834	Bahia	mother, calf	mother	F	300	Aug-Oct: all, Nov-: eod	5-10, 17-22
88719.09	17/09/2009	-13.952	-38.784	Bahia	five adults	adult	U	300	Aug-Oct: all, Nov-: eod	5-10, 17-22
87783.09	18/09/2009	-13.826	-38.767	Bahia	mother, calf, escort	mother	F	300	Aug-Oct: all, Nov-: eod	5-10, 17-22
88727.09	19/09/2009	-13.834	-38.717	Bahia	two adults	adult	F	300	Aug-Oct: all, Nov-: eod	5-10, 17-22
88720.09	24/09/2009	-13.850	-38.834	Bahia	four adults	adult	U	300	Aug-Oct: all, Nov-: eod	5-10, 17-22
81123.09*	24/09/2009	-13.942	-38.834	Bahia	four adults	adult	F	300	Aug-Oct: all, Nov-: eod	5-10, 17-22
81125.09*	24/09/2009	-14.004	-38.834	Bahia	four adults	adult	M	300	Aug-Oct: all, Nov-: eod	5-10, 17-22
81124.09*	25/09/2009	-13.897	-38.767	Bahia	mother, calf, four adults	adult	M	300	Aug-Oct: all, Nov-: eod	5-10, 17-22
81126.09*	25/09/2009	-13.887	-38.750	Bahia	mother, calf, four adults	mother	F	300	Aug-Oct: all, Nov-: eod	5-10, 17-22
81122.09*	27/09/2009	-13.887	-38.765	Bahia	mother, calf, three adults	adult	F	300	Aug-Oct: all, Nov-: eod	5-10, 17-22
87773.10	18/09/2010	-13.883	-38.843	Bahia	mother, calf, escort	escort	M	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
96380.10	23/09/2010	-13.835	-38.861	Bahia	mother, calf, escort	mother	F	300	Sep-Nov: all; Dec-: eod	2-8, 14-20

88726.10	28/09/2010	-13.883	-38.818	Bahia	mother, calf, escort	mother	F	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
87781.10	01/10/2010	-13.869	-38.807	Bahia	mother, calf	mother	F	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
88724.10	01/10/2010	-13.857	-38.792	Bahia	mother, calf	mother	F	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
84497.10	01/10/2010	-13.752	-38.786	Bahia	three adults	adult	F	500	Sep-Nov: all; Dec-: eod	0-23
87759.10	02/10/2010	-13.666	-38.846	Bahia	single	adult	F	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
87631.10	05/10/2010	-13.904	-38.889	Bahia	mother, calf	mother	F	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
87777.10	05/10/2010	-14.019	-38.879	Bahia	mother, calf	mother	F	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
84496.10	06/10/2010	-13.906	-38.871	Bahia	mother, calf	mother	F	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
87624.10	06/10/2010	-13.996	-38.855	Bahia	five adults	adult	U	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
87778.10	06/10/2010	-14.056	-38.838	Bahia	five adults	adult	U	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
102211.10	06/10/2010	-14.089	-38.815	Bahia	seven adults	adult	U	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
87776.11	29/09/2011	-13.858	-38.886	Bahia	mother, calf, escort	mother	F	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
87773.11	29/09/2011	-13.926	-38.870	Bahia	mother, calf	mother	F	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
87774.11	30/09/2011	-14.079	-38.827	Bahia	mother, calf, escort	escort	M	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
87769.11	01/10/2011	-13.877	-38.826	Bahia	mother, calf, two adults	mother	F	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
111869.11	07/10/2011	-13.722	-38.830	Bahia	single	adult	U	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
111868.11	10/10/2011	-14.086	-38.861	Bahia	mother, calf, two adults	escort	M	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
87783.11	10/10/2011	-14.175	-38.868	Bahia	mother, calf, two adults	mother	F	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
87768.11	22/10/2011	-13.920	-38.815	Bahia	mother, calf, escort	mother	F	300	Sep-Nov: all; Dec-: eod	2-8, 14-20
112694.12	08/08/2012	-21.837	-40.340	Rio de Janeiro	four adults	adult	U	400	all	5-10, 17-22
112717.12	08/08/2012	-21.785	-40.286	Rio de Janeiro	four adults	adult	U	400	all	5-10, 17-22
112702.12	20/10/2012	-18.085	-39.006	Abrolhos	eight adults	escort	M	400	all	5-10, 17-22
121189.12	20/10/2012	-18.096	-39.044	Abrolhos	pair	adult	M	400	all	5-10, 17-22
121194.12	20/10/2012	-18.061	-39.067	Abrolhos	eight adults	escort	M	400	all	5-10, 17-22
121195.12	20/10/2012	-18.051	-39.114	Abrolhos	mother, calf	mother	F	400	all	5-10, 17-22
112714.12	25/10/2012	-17.970	-39.053	Abrolhos	mother, calf, escort	escort	M	400	all	5-10, 17-22
120947.12	25/10/2012	-17.980	-39.133	Abrolhos	single	adult	M	400	all	5-10, 17-22
121193.12	25/10/2012	-17.969	-39.091	Abrolhos	mother, calf, escort	mother	F	400	all	5-10, 17-22

87632.12	26/10/2012	-18.104	-39.127	Abrolhos	mother, calf, escort	mother	F	400	all	5-10, 17-22
112712.12	26/10/2012	-18.127	-39.137	Abrolhos	mother, calf, escort	escort	M	400	all	5-10, 17-22
121192.12	27/10/2012	-18.089	-39.223	Abrolhos	four adults	adult	M	400	all	5-10, 17-22
121196.12	27/10/2012	-18.068	-39.222	Abrolhos	three adults	adult	M	400	all	5-10, 17-22
87775.12	28/10/2012	-18.009	-39.021	Abrolhos	mother, calf, escort	mother	F	400	all	5-10, 17-22
88480.12	02/11/2012	-18.031	-39.027	Abrolhos	mother, calf	mother	F	400	all	5-10, 17-22
111871.12	02/11/2012	-18.004	-39.040	Abrolhos	mother, calf	mother	F	600	eod	0-23

*Tagged with LIMPET tags.

The tagging team, comprising a driver, a photographer, one researcher responsible for biopsy sampling and another for tagging, operated from a tagging boat that was accompanied by a supporting vessel, usually a powerboat, in all years. During the 2012 cruise only, a second powerboat was not available and the research ship provided support during tagging operations instead. Tagging happened only in good conditions, calm sea and light/moderate winds (i.e., Beaufort scale less than 4).

Implantable (n = 108) and LIMPET (n = 5) (Low Impact Minimum Percutaneous Electronic Transmitter) tags from Wildlife Computers (Redmond, WA, USA) were used. Implantable tags were attached to the animals using a carbon fibre pole or with an Air Rocket Transmitter System (ARTS; Heide-Jørgensen et al., 2001); LIMPET tags were attached using a crossbow (Andrews et al., 2008). The targeted tag location on the animal was the base of the dorsal fin, because this area is most frequently exposed when animals surface to breath, maximising the chances of transmitter-satellite signal linking. The tagging boat approached target animals to distances between 3 and 10 m. When the pole was used, a maximum distance of 5 m was required. Photographs of the tagged animal's fluke, dorsal fin and fixed tag location on the animal's body were taken.

Data from tagged whales were obtained via the Argos system (Argos, 2016), which provides estimated locations and location quality information. Some tags were programmed to transmit in varying programmed duty cycles (e.g., transmitting every second/other day, every four days, etc.; Table 3.1), to maximize the potential period over which animals could be monitored.

Tagging of animals were conducted under permits issued by the Brazilian Institute of the Environment and Renewable Natural Resources (IBAMA, permit #009/02/CMA/IBAMA; process #02001.000085/02-27; ICMBio #11523–1).

3.2.2 Telemetry data processing

The timing of the location information received was irregular and, given location uncertainty, data were processed to derive continuous tracks from observed locations, from which regular estimated locations could be derived for analysis. This was achieved through a hierarchical form of the first difference correlated random walk (DCRW) model of Jonsen (2016), which estimates parameters for each individual in the sample.

Before tracks could be derived, the location data were filtered and formatted, as follows. Because of time resolution (nearest minute), some estimated locations from the same whale were recorded as being at exactly the same time, which would difficult the fitting of DCRWs. In those cases, R package *trip* (version 1.5; Sumner, 2016) was used to add one second to duplicate times, providing every location with a unique time value. The quality of locations that had a second added were compared to their original duplicates, and when either was recorded as very low quality, as informed by Argos (Argos, 2016) (i.e., quality B or A) and the other had a better quality (i.e., 0, 1, 2 or 3), the one with lower quality was excluded. A speed filter was also applied to the data, implemented with package *trip*, assuming a maximum travel speed of 12 km h^{-1} for humpback whales (Garrigue et al., 2010). To reduce the occurrence of long gaps between locations, for which no information was available, if a gap was longer than 10 days the derived track was split into two separate tracks.

After filtering the data, the hierarchical DCRW was fitted with R package *bsam* (Version 1.1.2; Jonsen, 2016). Package *bsam* derives location estimates from animal tracking data in user-defined regular time intervals, estimating parameters in a Bayesian state-space framework and using Markov chain Monte Carlo (MCMC) for inference. Starting values were automatically selected by *bsam* package. To estimate model parameters, 10,000 samples from two Markov chains were generated as a burn-in, then every 10^{th} of 30,000 samples were retained (3,000 samples per location) to estimate two locations per day (i.e., time step of 12 hours). Default *bsam* model output plots and plots of estimated locations *versus* original locations were inspected to check both model convergence and whether derived locations were roughly aligned with observed locations. Data from individuals for which the model did not converge of for which location estimates were clearly wrong, were discarded from the analysis. Model outputs from package *bsam* (e.g., posterior densities and model convergence plots) indicated poor model fit and/or convergence for tracks with less than 12 locations or less than five days of tracking. Tracks with less than 12 locations or less than 5 consecutive days of monitoring were therefore excluded from the analysis. A summary of the data used in the analysis is presented in Table 3.2. Appendix 3.1 presents the duration of tracks, the total observed locations per track, the

median, minimum and maximum per day for each track that was re-estimated using DCRW.

Because the objectives here were related to investigating distribution in the breeding area and to compare results to a previous study, derived locations beyond the limits of the survey area, in offshore waters or during migration, for example, were not considered in the analysis (Figure 3.2).

Table 3.2. Summary of the data set (presences only) used for modelling distribution. (Max locations = maximum number of locations per track within a year).

Year	Tracks	Locations	Initial date	Ending date	Max locations
2003	8	340	18-Oct	25-Dec	118
2005	11	359	11-Oct	27-Nov	83
2006	3	75	12-Oct	10-Nov	37
2007	7	281	08-Sep	18-Oct	78
2008	17	598	27-Aug	03-Nov	97
2009	9	399	14-Sep	09-Nov	98
2010	10	282	18-Sep	06-Nov	56
2011	8	314	29-Sep	20-Nov	63
2012	13	407	08-Aug	17-Dec	85
2013*	1	16	20-Aug	27-Aug	16
Total/Overall	87	3071	08-Aug	25-Dec	118

*locations from a whale tagged in 2012, which was monitored for more than 300 days.

3.2.3 Covariates

Candidate covariates considered in the spatial models were the same as in the line transect modelling study (Chapter 2). Time-varying variables had a different time range because the telemetry data were from 2003 to 2013 but the line transect data were available only for 2008 and 2012. Current speed close to the surface (Curr.sp) data were extracted from the *OSCAR Third Degree Sea Surface Velocity* dataset (ESR, 2009), with resolution 0.33×0.33^{o} (latitude × longitude) in 5-day intervals. Daily values for sea surface temperature (SST), with resolution 0.01×0.01^{o}, were extracted from *JPL MUR SST project* dataset (JPL MUR MEaSUREs Project, 2010). Daily wind speed at the surface of the sea (Wind.sp) values were extracted from the *Era-Interim* dataset (Dee et al., 2011), with horizontal resolution 0.125×0.125^{o}. Depth values were extracted form ETOPO1

(Amante and Eakins, 2009), for which the resolution is 0.1 × 0.1º. Slope was derived from ETOPO1, and was therefore at the same resolution. Distances from coast (Dist.coast) and from the shelf break (Dist.shelf), represented here by the 500 meters isobath, were measured with the *gDistance* function, *rgeos* R package (version 0.3-26; Bivand and Rundel, 2017). A factor to represent shelter was created by combining values of SST and wind speed in six classes, defined by quantiles of these covariates and following the same procedure described in the line transect modelling study (see Chapter 2). The logarithmic function of depth was used to aid model fitting for that covariate. A summary of covariates and other details is given in Table 2.1 (Chapter 2).

3.2.4 Data analysis

Interpolated locations were considered as presences in spatial models and, to provide contrasting data on absences, a set of pseudo-absences was created to represent were animals could have been, assuming that the survey area was accessible to all tagged animals. For every presence, five pseudo-absences were randomly created within the survey area. To select this ratio (five pseudo-absences per presence), different ratios (1:1, 3:1, 5:1 and 10:1) were tested by comparing outputs of the full model fitted to the different combinations of presences and pseudo-absences. The model fitted to the data with five pseudo-absences per presence produced the same output as with ten, but different than with small ratios (1:1 and 3:1), which indicates that five pseudo-absences per presence is sufficient to represent the underlying background of the inference area and that more pseudo-absences would not improve the modelling. Covariates were checked for correlation and collinearity and those which were strongly correlated (> 0.7) or had high (> 10; Hair et al., 2014) variance inflation factor scores (*vif* function, *car* R package) when together in a model, were not included in the same model at the same time.

3.2.4.1 Model fitting

As in the modelling of the line transect data (Chapter 2), two objectives guided the inclusion of candidate covariates in the models. First, to investigate habitat use, the initial habitat use model (HUM) included all available covariates except geographic position (latitude and longitude). In the distribution model (DIM) all covariates were considered to select the best predictors of distribution among the available covariates. The main

objective related to the DIM was to compare to compare to the AEM (Abundance Estimation Model) of Chapter 2.

Binomial models with logit link function were fitted in a modelling framework combining Generalised Estimating Equations (GEE; Hardin and Hilbe, 2002) and Generalised Additive Models (GAMs; Wood, 2006), using software R (R core team, 2017), and adapting approaches described in previously published studies (Pirotta et al., 2011; Jones et al., 2016; Russell et al., 2016). First, the SALSA (spatially adaptive local smoothing algorithm) method (Walker et al., 2011) was implemented with the *MRSea* R package (version 1.0.beta; Scott-Hayward et al., 2017) to automatically select the number and locations of knots for smooth (b-splines) functions of covariates in the models. The maximum number of knots were 8 for one-dimensional smoothers (i.e., all continuous covariates except geographic position) and 20 for two-dimensional smoothers (i.e., for geographic position), to prevent overly-wiggly smooths being fitted, which would difficult biological interpretation (see Wood [2006] for guidance on setting the number of knots). The SALSA method presents some advantages over other procedures, since it potentially improves the model fitting process with automatic selection of number of knots (ups to the limit imposed) and location of knots for smooth functions (Scott-Hayward et al., 2017).

The data were organised in correlation panels, with a panel for each set of presences within a track (one panel per track) and a different panel for each pseudo-absence (one panel per pseudo-absence). Using this panel structure was congruent with the assumption that locations within a tack were correlated, but that locations in different tracks were not, and that pseudo-absences were mutually independent. To account for the imbalance between the number of presences and pseudo-absences, pseudo-absences were given $1/5^{th}$ the weight of presences. At this stage, smoothed covariates for which no knot was indicated as significant, at $\alpha = 0.05$, as indicated by the robust standard errors (default output of MRSea package), were not considered in subsequent steps. Resulting models were refitted as GEEs with *geeglm* function (*geepack* R package, version 1.2-1; Højsgaardet al., 2006), to accommodate the residual correlation. The same spline smoothers were used. The QICu score (Pan, 2001), calculated using a custom formula written by Daniel J. Hocking (https://danieljhocking.wordpress.com/), was used for

model selection. The QICu score is adequate for selection of models fitted in a non-likelihood-based framework, such as GEEs, and is similar to AIC in the sense that it measures model fit and adds penalties based on model structure complexity (Pan, 2001). The significance of covariates was then verified using function *getPvalues* from *MRSea* R package (Scott-Hayward et al., 2017). This function fits several ANOVAs, with each covariate being the last to be included in the calculation, so that the p-values in the output are marginal.

In summary, model selection was conducted following three main criteria and in this order: 1) the full model was fitted with SALSA, and covariates for which no knot was estimated to be significant were not considered in subsequent steps; 2) a series of GEE-GAMs were fitted, leaving one covariate out at a time (backward step), and models' QICu were compared to verify if some of the covariates should be discarded; and 3) the marginal p-values for the remaining covariates in the GEE-GAMs were estimated and non-significant covariates were dropped. A conventional significance level was considered in all steps of covariate selection ($\alpha = 0.05$).

For the selected model, model performance was verified with Receiver Operating Characteristic (ROC) curves and confusion matrices, as per Pirotta et al. (2011), using R package *ROCR* (version 1.0-7; Sing et al., 2009). The ROC and confusion matrix can be used to calculate percentages of false positives and false negatives expected for the model, by comparing the predicted values to the observed. For comparison, the prediction grids containing covariate values from 2008 and 2012 used in the line transect modelling study (Chapter 2) were adapted to account for changes in the survey area (Figure 3.2) and used for predictions using the two selected models (HUM and DIM). The resolution of the prediction grid cells remained the same (i.e., 8 × 8 km). The contribution of each covariate in the final HUM was visualised with partial plots, with confidence intervals based on the GEE estimated uncertainty, adapting the R code provided in the supplementary material of Pirotta et al. (2011).

3.3 RESULTS

Data from 113 tags, deployed from 2003 to 2012, were available for the analysis. After filtering (see *Methods*, *Data preparation*), interpolated locations from 87 tracks were

available for investigating habitat use Figure 3.2, Table 3.1 and Table 3.2): comprising 62 tracks from females, 19 for males and 6 for animals of unidentified sex. Because five pseudo-absences were created per each presence (i.e., interpolated locations within the survey area; n = 3,071), 15,355 pseudo-absences were used, totalling 18,426 locations to be modelled.

Models performed well, with confusion matrices indicating 66% of correct predictions for the HUM, and 63% for the DIM, and the area under the ROC curve (AUC) was 0.691 and 0.732, respectively (Table 3.3).

Table 3.3. Performance of models fitted to telemetry data. HUM = habitat use model, DIM = distribution model, AUC = area under the ROC curve.

Model		**HUM**		**DIM**	
Correctly predicted		**66.2%**		**63.1%**	
AUC		**0.691**		**0.732**	
Confusion matrices		**Observed**		**Observed**	
	Predicted	**1**	**0**	**1**	**0**
	1	1871	5022	2330	6063
	0	1200	10333	741	9292
Total observed		3071	15355	3071	15355
		1	**0**	**1**	**0**
Percent of observed	**1**	60.9%	32.7%	75.9%	24.0%
	0	39.1%	67.3%	39.5%	60.5%

The final HUM model included smooth terms for current speed, SST, wind speed, coast distance and distance to the shelf break, and the factor variable shelter (Table 3.4 and Table 3.5). As a result using GEEs for dealing with auto-correlation in the data, confidence intervals for fitted relationships between the response variable and the covariates are very wide, except for SST (Figure 3.3). Depth had no significant knots when the full model was fitted in the first step of covariate selection (see section 3.2.4.1) and was not considered in subsequent steps. The fitted relationship for SST covariate presented a clear peak around 25ºC, for which the predicted probability of occurrence of whales was higher. A map of the probability of whale occurrence surface is presented in Figure 3.4.

Table 3.4. Marginal p-values for covariates in the final habitat use model (HUM).

Variable	**p-value**
Shelter	0.00659
Curr.sp	0.000164
SST	<0.0001
Wind.sp	0.0264
Coast.dist	0.0246
Shelf.dist	<0.0001

Table 3.5. Final habitat use model (HUM) coefficients. (co = cold, mo = moderate, wa = warm, li = light, Curr.sp = current speed, Wind.sp = wind speed, Coast.dist = distance to the coast).

Parameter	**Estimate**	**Std. err**	**Wald**	**P-value**	**Sign (α = 0.05)**
Intercept	-9.173	1.985	21.345	< 0.001	*
Shelter.co.mo	0.364	0.151	5.824	0.016	*
Shelter.co.st	0.380	0.171	4.926	0.026	*
Shelter.wa.li	0.011	0.172	0.004	0.951	
Shelter.wa.mo	0.010	0.219	0.002	0.963	
Shelter.wa.st	0.279	0.219	1.614	0.204	
bs(Curr.sp)1	-0.749	0.331	5.112	0.024	*
bs(Curr.sp)2	-0.552	0.259	4.522	0.033	*
bs(Curr.sp)3	-1.377	0.320	18.554	< 0.001	*
bs(Curr.sp)4	-0.638	0.326	3.829	0.050	
bs(Curr.sp)5	-1.930	0.685	7.926	0.005	*
bs(Curr.sp)6	0.554	1.266	0.192	0.662	
bs(SST)1	9.770	2.509	15.159	< 0.001	*
bs(SST)2	8.440	1.812	21.686	< 0.001	*
bs(SST)3	5.888	3.437	2.935	0.087	
bs(Wind.sp)1	0.719	0.359	4.010	0.045	*
bs(Wind.sp)2	-0.063	0.238	0.069	0.793	
bs(Wind.sp)3	0.428	0.391	1.198	0.274	
bs(Wind.sp)4	0.950	0.776	1.498	0.221	
bs(Coast.dist)1	0.834	0.429	3.777	0.052	
bs(Coast.dist)2	0.134	0.280	0.229	0.632	
bs(Coast.dist)3	1.508	0.547	7.594	0.006	*
bs(Coast.dist)4	-1.754	0.824	4.536	0.033	*
bs(Shelf.dist)1	0.675	0.277	5.950	0.015	*
bs(Shelf.dist)2	0.237	0.212	1.254	0.263	
bs(Shelf.dist)3	1.228	0.232	28.095	< 0.001	*
bs(Shelf.dist)4	1.379	0.266	26.810	< 0.001	*
bs(Shelf.dist)5	1.999	0.390	26.214	< 0.001	*
bs(Shelf.dist)6	0.147	0.581	0.064	0.801	

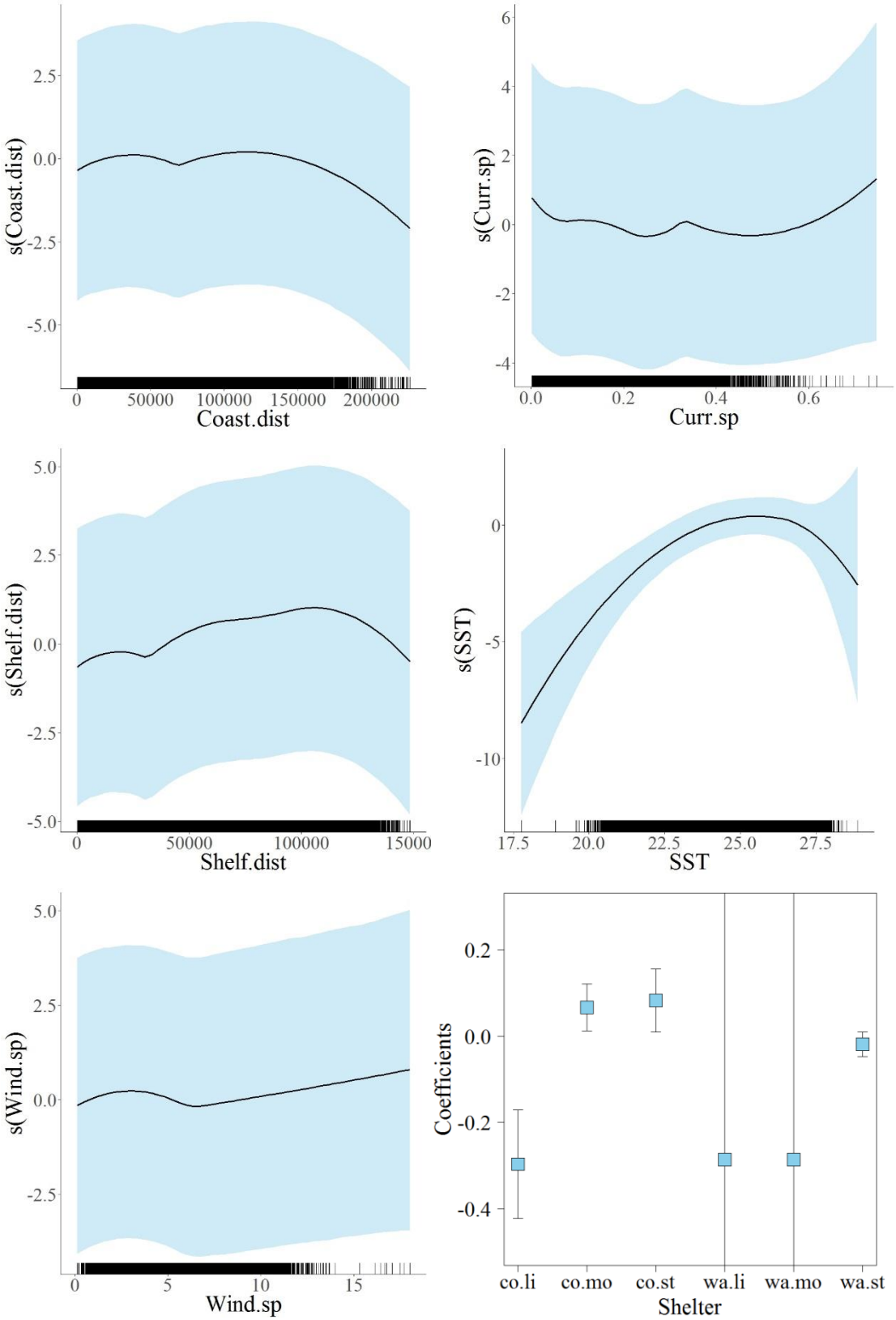


Figure 3.3. Fitted relationships for smooth functions of covariates in the final HUM. Error bars in the "Shelter" plot represent 95% normal confidence intervals. (co = cold, wa = warm, li = light, mo = moderate, st = strong).

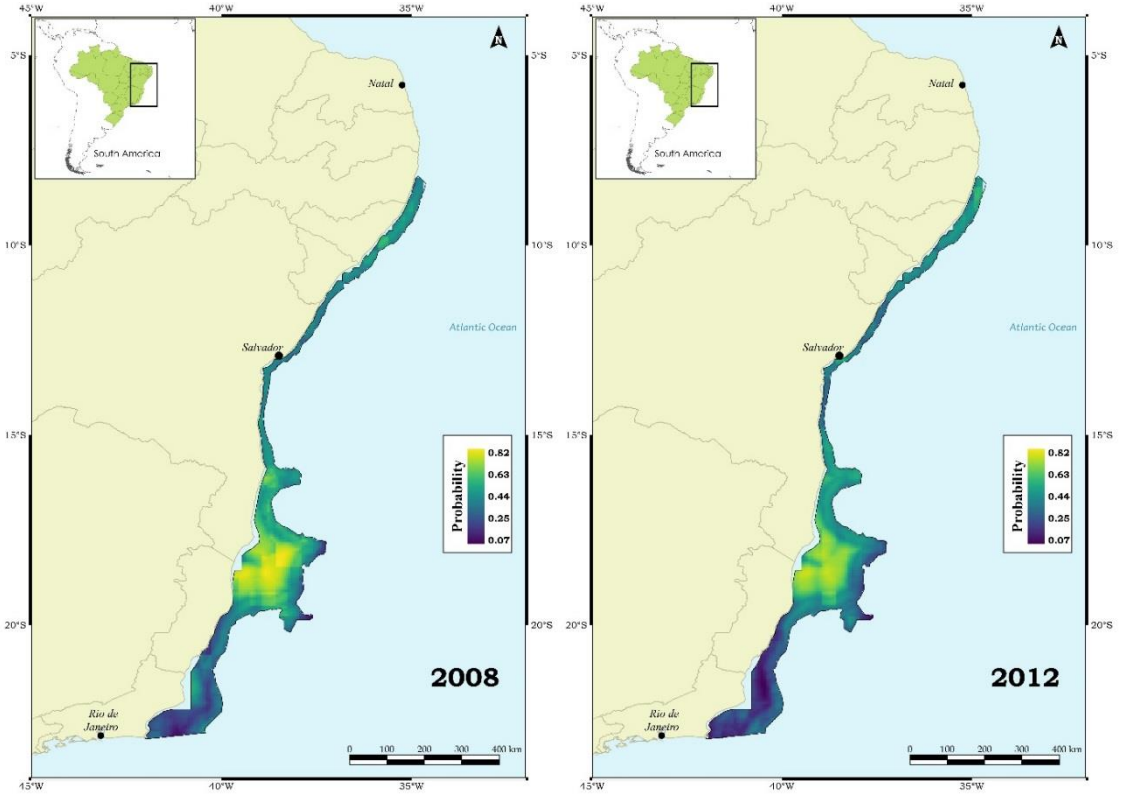


Figure 3.4. Occurrence probability surface, predicted using the habitat use model (HUM).

The final DIM differed from the HUM in terms of covariates, with latitude and longitude considered instead of SST (pre-model fitting, i.e., in the full model). Shelter and wind had no significant knots at the first step of covariate selection, and depth had a non-significant marginal p-value in the last step of covariate selection for the DIM (Table 3.6 and Table 3.7). Higher probabilities of encountering whales were predicted for the region around the Abrolhos archipelago and to the south (Figure 3.5). Maps resulting from the method investigated here and from Chapter 2 are compared in Figure 3.5. A summary of covariates retained in the final HUM and DIM is presented in Table 3.8.

Table 3.6. Marginal p-values for covariates in the final distribution model (DIM).

Variable	**p-value**
Curr.sp	0.00668
Coast.dist	<0.0001
Shelf.dist	<0.0001
Latitude and longitude	<0.0001

Table 3.7. Distribution model (DIM) results.

Parameter	Estimate	Std. err	Wald	P-value	Sign ($\alpha = 0.05$)
Intercept	-5.605	0.674	69.185	< 0.001	*
bs(Curr.sp)1	-0.547	0.335	2.660	0.103	
bs(Curr.sp)2	-0.643	0.262	6.005	0.014	*
bs(Curr.sp)3	-0.570	0.290	3.856	0.050	*
bs(Curr.sp)4	-1.100	0.836	1.731	0.188	
bs(Curr.sp)5	-1.020	0.287	12.644	< 0.001	*
bs(Curr.sp)6	-0.402	0.484	0.690	0.406	
bs(Curr.sp)7	-1.093	0.928	1.385	0.239	
bs(Coast.dist)1	1.434	0.288	24.701	< 0.001	*
bs(Coast.dist)2	0.512	0.296	2.994	0.084	
bs(Coast.dist)3	0.042	0.527	0.006	0.937	
bs(Coast.dist)4	-5.109	1.293	15.615	< 0.001	*
bs(Shelf.dist)1	0.702	0.233	9.061	0.003	*
bs(Shelf.dist)2	0.643	0.183	12.304	< 0.001	*
bs(Shelf.dist)3	3.099	0.379	66.943	< 0.001	*
bs(Shelf.dist)4	1.748	0.455	14.773	< 0.001	*
XY.1	12.740	1.847	47.568	< 0.001	*
XY.2	2.689	0.305	77.781	< 0.001	*
XY.3	4.599	0.819	31.563	< 0.001	*
XY.4	-7.539	1.353	31.072	< 0.001	*
XY.5	2.771	0.708	15.329	< 0.001	*

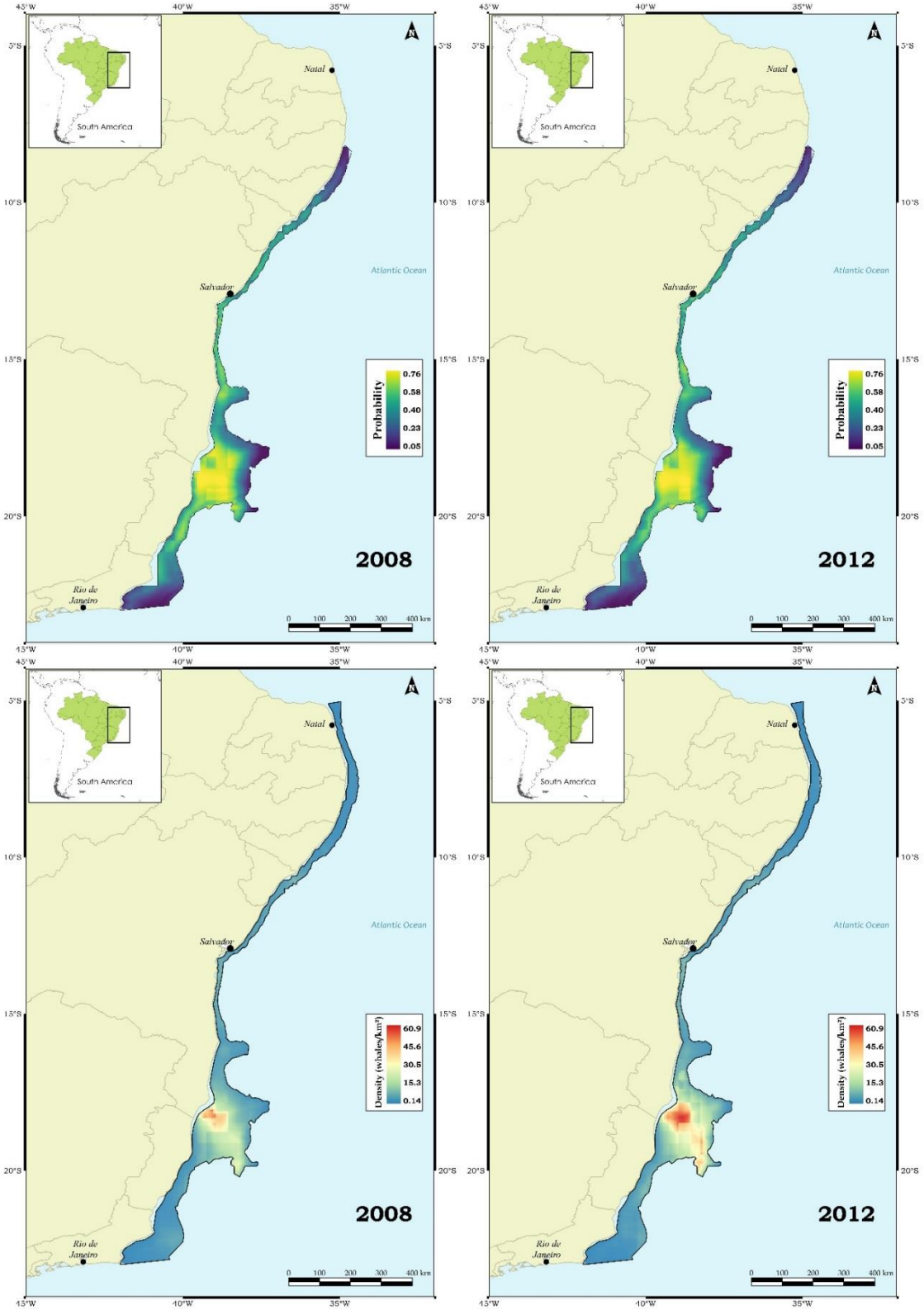


Figure 3.5. Upper panel: Occurrence probability surfaces, predicted using the DIM; Lower panel: AEM maps (Figure 2.10, Chapter 2), showing density surfaces, for comparison.

Table 3.8. Covariates retained in each model fitted to telemetry data and to line transect data. HUM = habitat use model, DIM = distribution model, AEM = abundance estimation model.

	Telemetry data		Line transect data	
Covariate	**HUM**	**DIM**	**HUM**	**AEM**
s(Curr.sp)	*	*	*	*
s(Depth)				
s(Dist.shelf)	*	*	*	*
s(Dist.coast)	*	*	*	*
Shelter	*		*	*
s(Slope)				
s(SST)	*		*	
s(Wind.sp)	*			*
s(lat,lon)	—	*	—	*

3.4 DISCUSSION

In this chapter, the distribution and habitat use of humpback whales off Brazil were investigated from spatial models applied to locations obtained through animal tracking and covariates. Issues that such data naturally present, such high spatiotemporal correlation in locations, were taken into account in the analysis methods. Although higher uncertainty in fitted covariate-response relationships was found here, and consequently in habitat use inferences, distribution patterns agree with those from Chapter 2, where spatial models were applied to line transect data.

Despite differences in the nature of the data (individual vs. space, respectively) and statistical tools used, some of the outputs from modelling locations derived from telemetry data as presences in a presence/pseudo-absence modelling approach support findings from distribution models fitted to line transect data (Chapter 2). Despite the higher uncertainty in most of the present fitted covariate-response relations (Figure 3.3), SST showed a clear peak around the same range of temperatures in both the present study and in the line transect modelling (Figure 3.6), which probably reflects how strongly the habitat use for these animals is related to temperature in the breeding area. The apparent agreement between the two methods is not only reassuring for what is known about habitat use for this population, but also shows that similar conclusions can be drawn on this regardless of which of the two methods is used.

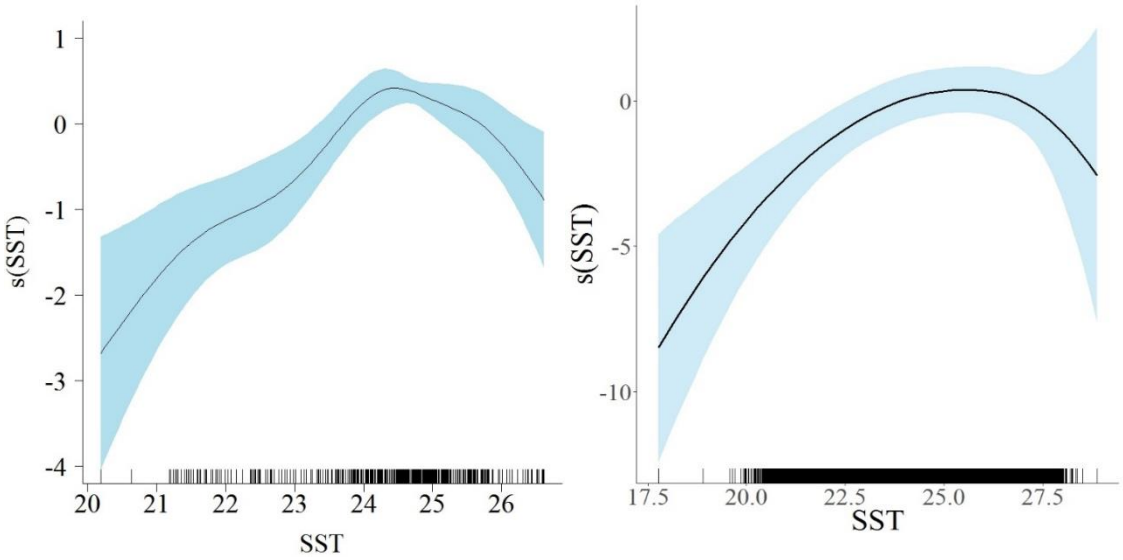


Figure 3.6. Comparison of smooth functions fitted to SST in the habitat use models, from line transect data (left) and satellite telemetry data (right).

Although the present results are reassuring, they must be interpreted with caution. For example, by creating pseudo-absences from random locations in the entire area, it was assumed that the entire survey area to be available to the animals at any time. However, this is an arbitrary decision and the available area from the animals' perspective could be very different (Aarts et al., 2008). For example, an animal tagged in the vicinity of the Abrolhos archipelago (Figure 3.1 and Figure 3.2) would take at least two days to reach the southernmost portion of the survey area, swimming in a straight line at 12 km h^{-1} (i.e., the maximum travel speed [Garrigue et al., 2010] used as a criteria in the data pre-processing/filtering). The restriction of the area of inference (survey area) in this chapter (compare north extents in upper and lower panels of Figure 3.5) was adopted to reduce the effects of violating the area availability assumption (Aarts et al., 2008). Further restrictions could be investigated at the expense of inferring over a smaller region and considering fewer data. Another possibility to avoid issues from non-uniformly distributed tagging locations could be to truncate the first days from each track. This was not possible here because many tracks presented relatively short periods of locations within the survey area.

The extent of the area of inference also has a major influence on where created pseudo-absences are placed, and in here they served to represent the background

environment where animals were not observed. The restriction of the area considered in this study potentially also minimized the effects related to this issue, since areas further north from the limits of the survey area (Figure 3.2) were visited by only two animals. Also, because of imbalance in the tagging locations and numbers of animals tagged across years (Table 3.2), and because as result data from all years were pooled to model distribution, it was impossible to evaluate temporal variation in distribution and habitat use patterns. Data was assumed to be representative for the population as a whole and for the period between 2003 to 2013. Year was not considered as a covariate in the analysis and inferences presented here must be interpreted as the distribution pattern for the population overall. Investigating temporal variation in distribution could help understanding the potential expansion of population range in the breeding area (Pavanato et al., 2018), but spatio-temporally balanced data within the area of inference is needed for that. One possible option is to restrict the area of inference in ways that allow meeting the above criteria, at the expense, again, of inferring over a much reduced area.

The criteria used here for selecting covariates retained in the final models are a combination of different approaches to look for a substantive contribution from covariates to model fitting (through the *QICu* score) and for enough information in the fitted relationship (through p-values). The reason for the above criteria to differ from that adopted in Chapter 2 is that a likelihood cannot be calculated for GEEs, for which *QICu* is the most indicated score for GEE model selection (Pan, 2001). The covariate inclusion in models was also decided based on significance (at $\alpha = 0.05$) because some covariates presented high uncertainty on the fitted relationships (Figure 3.3), which could not be precisely estimated. This criteria was used to exclude covariates that, although indicated by *QICu* it could improve the model, the relation to the response was uncertain. Despite the criteria having the potential to exclude important covariates, models performed well (Table 3.3) indicating that distribution and habitat use inferences can be made from the present outputs.

The first step of covariate selection (i.e., observing robust standard errors from model fitting with SALSA) led to discarding of covariates that, even before the residual auto-correlation was accounted for, were not estimated precisely enough to be considered significant. Because serial autocorrelation may cause non-important covariates to appear

more significant (Aarts et al., 2008), GEEs permit more realistic estimation of uncertainty for covariate-response relations. However, fitted relations can be difficult to interpret (Pirotta et al., 2011), especially for those covariates with wide confidence intervals in their fitted relationships, as in Figure 3.3. The second step of covariate selection (observing the QICu scores) had been used in previous studies analysing similar data to the present and is considered a reasonable way to account for the residual auto-correlation issue (Pirotta et al., 2011; Jones et al., 2017). The last step of covariate selection (marginal p-values) was the final check for contribution of covariates to the models. The combination of the three criteria adopted improves the chances of only retaining important covariates related to animal distribution. Despite covariate selection criteria adopted, all covariates but SST presented large confidence intervals (Figure 3.3) and must be interpreted with caution, which results from accounting for the autocorrelation in the data, as pointed out in Pirotta et al. (2011). However, covariates selected can be considered related to occurrence as discussed above.

Similarly to the model used to investigate habitat use in Chapter 2, the initial HUM did not include latitude and longitude. This procedure was adopted because these covariates have no logical biological interpretation for habitat use, and also because they were both strongly correlated with SST. The DIM, similarly to the AEM in Chapter 2, included these covariates as potential explanatory terms, because its objective was to identify those features that describe most variability in whale occurrence and to generate the best predictive map for potentially informing management. In the present chapter, the DIM map was created to be compared the AEM map shown in Figure 2.10 (Figure 3.5). The present final HUM adds wind speed to the already identified important covariates related to habitat use (Martins et al., 2001; Bortolotto et al., 2017; Pavanato et al., 2018). However, the combination of SST and wind speed represented by shelter, had been already identified as very important in the line transect models. Using autoregressive models on line transect data, Pavanato et al. (2018) identified bathymetry and distance to the shore to be related to the occurrence of humpback whale groups in the area. Wind speed was also considered as a candidate covariate in that study, but it was not retained in that final model. In the present study, it is unclear what is the relation between wind and whale occurrence from the partial plot for that covariate (Figure 3.3) or by looking at

its estimated coefficients (Table 3.5). However, by looking at the fitted coefficients for shelter (Figure 3.3), wind seems to be important when the water is relatively colder, which agrees well with the results found from modelling line transect data. Other covariates in the final HUM include current speed and distance from the shelf and to the shore, which may be related to calf survival probabilities, protection against predators or in habitat selection for specific reproduction-related groups (Connor and Corkeron, 1999; Félix and Botero-Acosta, 2011). Because in this study, locations derived from modelled telemetry data were used to investigate the population distribution overall, sex and other individual characteristics were not incorporated in the models. The reason for not using this information was that there is no clear way to allocate sex and other individual characteristics to pseudo-absences within the GEEs framework. One option would be to fit models separately for males and females, since this information is available (Table 3.1). However, sex was unbalanced in the present data with the majority of animals being females. As discussed in Chapter 2, the distribution patterns found in that study and now supported in this, are consistent to what could be expected for population distribution driven by female habitat selection. As also noted in Chapter 2, movement models, with individual characteristics considered to estimate behaviour states, should provide important information on habitat use for these animals. Such models can be used to quantify the relationship between covariates and the probability of animals being in determined behaviour states (Roncon et al., 2018; Jonsen et al., 2019), therefore to investigate habitat use, not necessarily from a spatial perspective. In this sense, comparisons to line transect models as presented here would be more difficult.

Table 3.9. Comparison of the two spatial modelling methods considered here, of line transect data and satellite telemetry data (as presences and pseudo-absences).

Feature	Line-transect	Satellite telemetry
Sampling/data	Pro: spatial data are sampled from places visited by the observation platform, therefore controlled in the survey design Con: survey must cover a representative portion of the area	Pro: spatial data is sampled from places visited by the animals Con: no information about places not visited by the tracked animals
Modelling	Pro: potentially more precise outputs; more well consolidated model fitting procedures available Con: Data points along the same line are correlated in space and time.	Con: data points highly correlated in space and time within each individual; requires arbitrary number of pseudo-absences to be created; potentially large uncertainty in fitted smooth terms because of high spatiotemporal correlation (within GEEs framework).
Coverage/ Survey area	Pro: can be controlled in the survey design Con: limited temporally by the period of the survey and restricted spatially by survey platform limitations	Pro: animals may visit areas not assessable to survey platforms Con: limited temporally by tag longevity and restricted spatially by animal displacement restrictions

It is certainly important to investigate the potential effect of human activities on the occurrence of whales in the area (Pavanato et al., 2018). However, to evaluate the potential effects from human activities in the distribution of animals, either data from before the presence of such activities are needed, or drastic changes in the distribution or intensity of them must happen before its possible to identify a redistribution, coinciding with the change on activities. The fact that the two methods considered here agree in terms of distribution and habitat use inferences, means that both may be used in conjunction in informing conservation or management actions.

Modelling animal distribution using line transect data may have some important advantages in data collection, survey design and modelling techniques, such as being able to control the places visited in the survey design (Table 3.9). However, line transect surveys additionally allow to derive abundance estimates when the assumptions of

distance sampling are met (Buckland et al., 2015). If effort at field and in the analysis are made to meet the method assumptions, line transect data can provide information on both distribution and abundance.

On the other hand, satellite tagging of whales allow data to be collected remotely, meaning that the field work necessary for the method is restricted to that of tagging the animals. Also, the data may allow different approaches for investigating habitat use and distribution, because they include information about movement of individuals (Aarts and Matthiopoulos, 2010). Because close proximity to the animal is required for tagging, other useful information may also be collected. For example, it is common to collect skin and blubber biopsies from tagged whales, taking advantage of the tagging procedure, which can be used for determining sex, for genetic studies, for hormone-related and contaminant studies (Heide-Jørgensen et al., 2006; Reisinger et al., 2014). Tagging of whales is a much more invasive field procedure (Alves et al., 2010), but provides data that are impossible to obtain from line transect surveys only. The two methods compared here are complementary and may be used in conjunction to expand the spatiotemporal coverage of studies on distribution or habitat use, therefore providing better information to be used for implementing conservation and management actions when needed. For example, survey efforts on investigating distribution of a population can be split between tagging some animals in one portion of the area of interest and surveying another portion.

APPENDIX 3.1

Table A3.1. Locations for tracks (n = 87) used in hierarchical first difference correlated random walk (DCRW), implemented using *bsam*, for which derived tracks were used in the distribution analysis.

Track	Total	Duration	Median/day	Min./day	Max./day
20687.03	297	21	14	10	19
21791.03	170	13	13	7	18
21800.03	274	30	9	4	15
21809.03	50	16	3	0	7
21810.03	35	13	3	0	4
24640.03	22	8	2	1	5
24642.03	588	204	2	0	10
27259.03	50	38	1	0	6
7617.05	17	16	0.5	0	4
10946.05	118	56	2	0	8
24641.05	142	56	2	0	7
26712.05	29	16	2	0	4
27258.05	57	29	2	0	6
27259.05	25	18	1	0	3
27261.05	62	46	1	0	6
33000.05	17	25	0	0	4
33001.05	19	6	3.5	0	6
37231.05	18	18	1	0	2
37234.05	26	32	0	0	5
33001.06	52	15	3	2	8
37229.06	51	18	3	0	6
50682.06	22	16	1	0	5
27261.07	36	24	1	0	4
37286.07	35	14	3	1	4
37288.07	102	36	3	0	6
42521.07	72	40	2	0	4
42525.07	57	15	4	0	9
50686.07	72	15	4	2	9
50687.07	15	11	1	0	4
87759.08	75	11	6	5	10
87760.08	252	44	6	1	10
87761.08	189	30	6.5	2	9
87762.08	261	37	7	2	15
87763.08	278	54	5	0	13
87764.08	275	46	6	1	10
87765.08	244	32	8	2	12
87766.08	86	19	5	1	8
87767.08	26	5	5	1	9
87768.08	219	52	5	0	10
87769.08	80	53	1	0	7
87770.08	16	8	1.5	0	5
87773.08	252	41	6	2	10
87774.08	82	18	4	3	8

87775.08	247	35	7	3	12
87776.08	47	26	1	0	6
87778.08	122	34	3	0	11
81123.09	27	7	3	2	6
81125.09	21	7	3	2	5
81126.09	48	10	4.5	3	9
87771.09	383	60	6	1	12
87773.09	230	55	4	0	9
87774.09	233	49	5	0	9
87783.09_2	614	123	5	0	17
88719.09	84	48	2	0	4
88727.09	171	63	3	0	7
84497.10	333	27	12	3	21
87631.10	39	6	6.5	4	9
87773.10	28	12	2	0	7
87777.10	136	22	7	3	10
87778.10	117	21	6	0	10
87781.10	51	10	4	2	9
88724.10	49	8	5.5	5	11
88726.10	99	22	4	2	8
96380.10	92	19	4	3	12
102211.10	214	31	7	4	10
87768.11	272	29	9	6	13
87769.11	25	29	0	0	4
87773.11	205	35	6	0	10
87774.11_2	45	14	3	1	7
87776.11	21	22	1	0	6
87783.11	81	19	3	1	9
111868.11	86	18	5	0	8
111869.11	123	26	6	0	9
84480.12	310	45	7	0	15
87632.12	358	41	9	2	16
87775.12	15	12	1	0	4
111871.12	2492	169	13	1	30
112694.12	75	7	11	7	15
112702.12	172	32	6.5	0	12
112712.12	101	23	4	0	9
112714.12	48	7	6	2	15
112717.12	127	17	7	4	14
120947.12	84	9	10	6	12
121189.12	567	96	6	0	14
121189.12_6	49	8	6.5	2	9
121193.12	290	35	8	3	18
121195.12	52	7	6	5	10

CHAPTER 4

POTENTIAL IMPACT FROM OIL SPILLS TO HUMPBACK WHALES IN BRAZIL

ABSTRACT

With the continuing increase in human exploitation of natural resources in the marine environment, the prevalence of impacts on marine species is also expected to increase. The coast of Brazil is no exception, with oil production activities presenting a potential source of disturbance and contamination for cetaceans, including the humpback whale population that overwinters in the area. However, little is know about the risks such activities potentially pose to cetaceans present in the area. To evaluate the risk to which these whales could be exposed should oil spills from production fields in the vicinity occur, an oil dispersion simulation was investigated here. This exercise consisted of 1) simulating paths that oil hypothetically spilled from oil production fields would take, 2) calculating the density of simulated oil tracks across the area of interest and 3) combining simulated oil densities and whale densities to estimate risk across the area. The method proposed here assumes that oil is dispersed due to the current at the surface. Results suggest that humpback whales in the southern portion of their range in Brazilian waters are subject to a greater risk of impact from oil spills, due mainly to the high concentration of oil production fields in that area. Results presented here can help to inform decisions on management of oil production activities and mitigation of their impacts on humpback whales.

4.1 INTRODUCTION

No area in the marine environment is unaffected by human influence, with coastal regions suffering more from potential cumulative effects of multiple threats than oceanic environments. Among the various sources of impact from human activities are those related to the oil industry (Halpern et al., 2008). The physical presence of oil production structures in the marine environment are long known to impact its surroundings, sometimes causing drastic changes in the local ecosystems (Wolfson et al., 1979). Perhaps the most severe potential impacts from oil exploration activities in the marine

environment result from oil spills. Sources of oil spills include both fixed structures (e.g., pipes, drilling and oil production platforms), such as in the Deepwater Horizon accident (Wallace et al., 2017), or from ships transporting oil (Brekke et al., 2005). Oil spills can affect the marine and coastal environments in multiple ways, and several classes of marine life, from microbial communities (Kleindienst et al., 2015) to marine mammals (Dias et al., 2017), can be heavily impacted.

Cetaceans seem to be able to detect and avoid oil in the water, although this ability depends on oil thickness and colour, and they may still be vulnerable to contact with toxic oil components (Geraci et al., 1983; Smultea and Wursig, 1995). Killer whales have been observed swimming through heavily oiled waters in the wild, indicating that either the animals did not detect the oil or did not avoid it (Matkin et al., 2008). It has been suggested that, because of lack of fur (Helm et al., 2015), and characteristic tightness of intercellular bridges and thick epidermis (O'Hara and O'Shea, 2011), oil would not adhere to the body of cetaceans or directly penetrate it. However, animals exposed to oil may directly absorb its toxic components by ingestion, inhalation or through contact with the eyes and mouth mucosae (Geraci and St Aubin, 1990). Dias et al. (2017) documented the exposure of cetaceans to oil from the Deepwater Horizon spill, which happened in April 2010 in the Gulf of Mexico, U.S.A., and reported 85 occurrences of cetaceans swimming in or near petroleum products, or with oil adhered to their skin, until about two years after the incident.

Oil spills in Brazil represent an important potential threat to marine and terrestrial coastal species, and efforts to identify vulnerable species and priority areas for protection along the coast to inform management of operational responses to accidents have been made (Ruoppolo et al., 2017). Cetaceans were indicated to be priority species for protection in the event of oil spills. The largest, and perhaps the most memorable marine oil spill in Brazil, happened on 7 November 2011, when a pressure spike occurred during the drilling of an exploratory well at a depth of 1000 meters about 120 kilometres from the coastline of Rio de Janeiro State, in the Campos Basin (ANP, 2018). Between 83 and 140 m^3 of oil were estimated to be spilled, causing a 163 m^2 wide patch (ANP, 2018). Planning responses to oil spills is fundamental to protect the marine environment and the decision-making process must be very fast and effective to prevent as much impact as possible.

Usually managers and stakeholders responsible for the oil spill responses rely on information from environmental experts and tools to support decisions (Romero et al., 2013).

Humpback whales in Brazil concentrate mainly near the Abrolhos Bank (Chapters 2 and 3) and the population is recovering rapidly (Chapter 5). Even though new information on their distribution, abundance (Chapters 2 and 3; Bortolotto et al., 2017; Pavanato et al., 2017) and growth rate (Wedekin et al., 2017) recently became available, evaluations of potential impacts for this population in the area are not fully developed. The overlap of whale presence and human activities to identify areas at higher risk of impact was investigated in Martins et al. (2013), which included proposed areas for oil and gas exploration activities. Areas in the south of Abrolhos Bank were indicated to be at higher risk of impact. Castro et al. (2014) noted that marine protected areas in Brazil cover a very small portion of the areas mostly used by this population, implying that the level of protection provided by these areas is low.

In the second chapter of this thesis, results suggested that the distribution of whales was negatively correlated with the distribution of the oil fields. However, it is unclear whether oil spills could impact the habitat used by humpback whales during their breeding/calving season. Here a method to investigate the potential of oil spilled from production fields reaching these habitats was investigated.

The objective of this chapter was to develop a method to identify areas of higher risk of whales being directly impacted by spills from oil production fields. This method was applied to the humpback whale population that winters in Brazilian coastal waters. The method presented can be easily implemented and can be used to map and calculate the risk of oil spilled in the ocean to reach areas where animals are expected to be present. Outputs from this method should be useful to inform mitigation of the impact of marine oil exploration activities by identifying areas where oil exploration may represent a threat to this population. Strong assumptions about how oil is dispersed in the sea require caution in interpreting results, especially regarding the temporal scale to be considered. However, the method is a fast and easy to interpret first step to provide potentially useful information for risk assessment.

4.2 MATERIAL AND METHODS

The area considered for oil dispersion simulations comprised the typical area of distribution of humpback whales in the coast of Brazil (see Figure 1.1, Chapter 1) and nearby waters, to include oil fields in the vicinity (n = 99; Figure 4.1). The time period considered was around the expected annual peak of abundance of whales in the area, August–September (Martins et al., 2001; Andriolo et al., 2010).

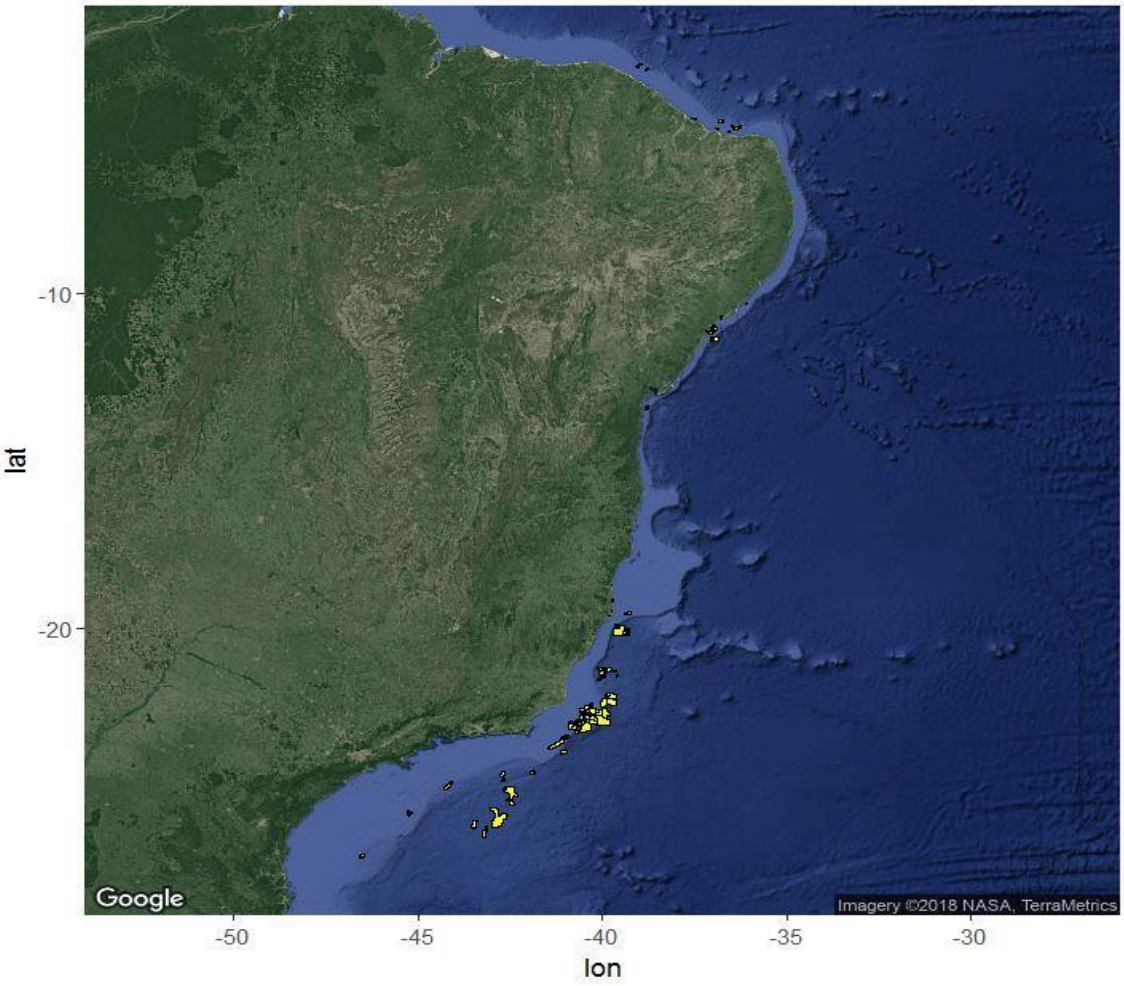


Figure 4.1. Oil fields (yellow polygons with black contour) considered in simulations as sources of oil dispersion (see Figure 2.2, Chapter 2, for more detail).

Oil field polygons, referred to here as "fields", were downloaded on 5 September 2018 from the Brazilian National Agency of Petroleum and Gas (Agência Nacional de Petróleo e Gás, ANP) online map (ANP, 2018). Fields considered here were those at production stage on 14 of August 2018. Current speed data were obtained from the Global Ocean Physical Reanalysis product (*GLOBAL REANALYSIS PHY 001 030*; Fernandez and

Lellouche, 2018), through the Copernicus marine environment monitoring service website (www.mercator-ocean.eu). This reanalysis uses wind speed close to the surface as described in Large and Yeager (2009) to estimate realistic values for oceanographic parameters (Fernandez and Lellouche, 2018). There is a large concentration of oil fields in the southern portion of the study area to the south of 20ºS (Figure 4.1). Most oil fields are located over the continental shelf or along the shelf-edge.

Simulation of oil dispersion was conducted starting from the centroid of each production oil field polygon at sea (Figure 4.1), or the centroid of the portion at sea in polygons that partially occurred inland. Polygon centroids were considered as the "source" of oil. Centroids within enclosed bays were excluded from the analysis.

The method investigated here followed three steps to simulate oil dispersion:

1. Starting from the source (e.g., a centroid of an oil production field polygon), the current speed at the sea surface was extracted for a given starting day;
2. Current speed values (zonal and meridional) were used to calculate the displacement of oil at the surface, as a straight line, assuming that the oil travels at the surface under the influence of the same current for an entire day. This resulted in a new location for the following day;
3. Current speed values were extracted for the new location and day, and step 2 was repeated a number of times, determined by the time length of the simulation.

The time length of a simulation and the effect of current speed determined how far the simulated oil was dispersed. Each simulation of oil dispersion is defined here as an oil track. For example, to investigate where the oil could be carried by the currents after two months of a spill, the time length for simulations was 60 days.

To investigate variation in oil tracks over a period within which an oil spill was simulated to happen, steps 1 to 3 above were repeated a number of times, starting from the same source (i.e., field centroid) but with each new simulated oil track starting one additional day after the previous one. The group of simulations from the same source and with different starting days provided pictures of the variation in oil tracks for the period considered. For example, to investigate oil tracks from spills happening in August 2012,

30 simulations starting on each day (except the last day) of August 2012 were run for each source (Figure 4.2). This period of time in which an oil spill was simulated to happen (from 1 August to 30 August 2012, in the above example) differs from the simulation's time length: the former represents how many starting days were considered for oil spill to happen and the latter represents how many steps of oil dispersion, of one day long each, were considered in each oil track. To demonstrate different features of the method, multiple scenarios were considered by varying the initial day and the time length. The period over which each scenario was assessed was fixed at 30 days (vertical axis in Figure 4.2) to facilitate comparison among scenarios. This means that potential spills within a month period were evaluated in every scenario.

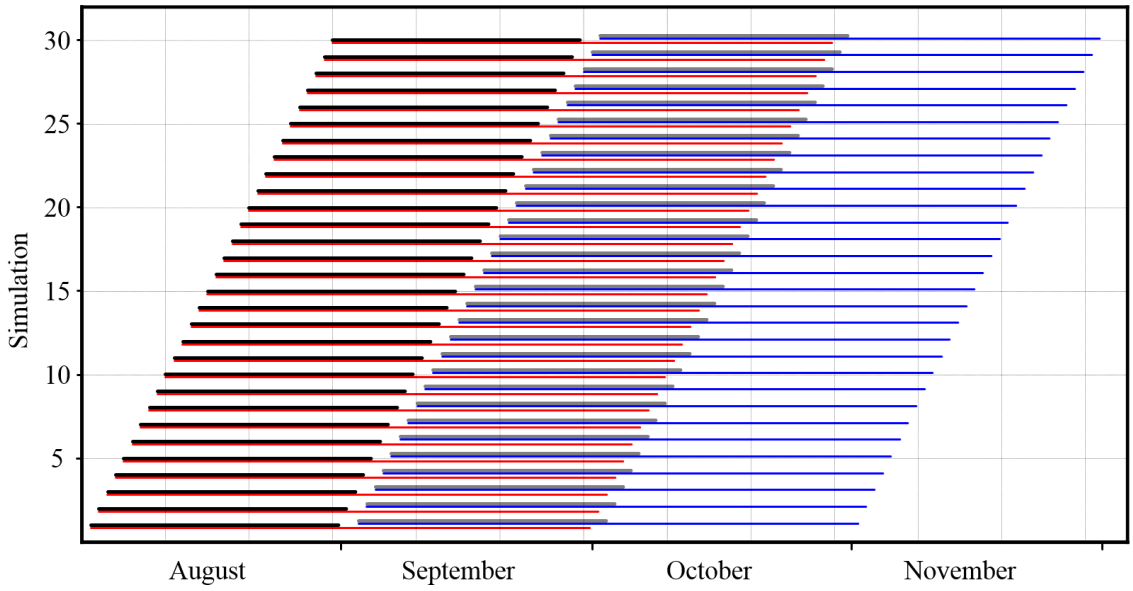


Figure 4.2. Schematic showing temporal range for simulations with 30 and 60 days of time length. Black and red lines represent simulations starting in August, with time lengths of 30 and 60 days, respectively. Grey and blue lines represent simulations starting in September, with time lengths of 30 and 60 days, respectively.

To illustrate how simulated oil tracks vary within the winter season and between years, initial days were set at 1 August and 1 September in 2008 and 2012, the years that information on whale distribution was available (Chapter 2 and 3). To illustrate the effect of varying time lengths, simulations were run for 15, 30 and 60 days. For 60 day simulations , for example, the simulated oil tracks which started on the last starting day

for September 2012 (30 September 2012) had its last simulated location on 30 November 2012, as indicated by the top blue line in Figure 4.2.

To illustrate inter-annual variability, simulations with a time length of 30 days were run for August (i.e., initial day = 1 August) 2012 to 2016 (the last year for which current speed data were available). Plots for presenting simulations were created with R package ggmap (Kahle and Wickham, 2013). The code used to simulate oil tracks is presented as a commented "rmarkdown" document (R package rmarkdown version 1.8; Allaire et al., 2017) in Appendix 4.1). Simulation scenarios considered here are summarised in Table 4.1.

Table 4.1. Summary of simulation scenarios.

Objective	**Time length**	**Period**
Illustrate variation in winter months, between 2008 and 2012 and between time lengths	15	August 2008
	30	August 2008
	60	August 2008
	15	September 2008
	30	September 2008
	60	September 2008
	15	August 2012
	30	August 2012*
	60	August 2012
	15	September 2012
	30	September 2012
	60	September 2012
Illustrate inter-annual variation	30	August 2012*
	30	August 2013
	30	August 2014
	30	August 2015
	30	August 2016

*Used in both objectives and for risk assessment.

4.2.1 Risk assessment

To illustrate the application of the oil dispersion simulation method for investigating risk of impact to whales, a 30-day simulation exercise with initial day equal to 1 September 2012 was conducted. The density surface map for 2012 from Chapter 2

(Figure 2.10) and the probability of occurrence surface map for the same year from Chapter 3 (Figure 3.3) were used to represent whale distribution.

With the 8×8 km prediction grids from Chapter 2 ("line transect grid") and Chapter 3 ("telemetry grid"), numbers of individual simulated oil tracks that crossed each grid cell were used to generate surfaces of oil track counts. Relative oil track counts, relative whale density and relative probability of occurrence were calculated for each grid cell, from the respective surface, by dividing grid cell values by the sum of values over the entire surface. Resulting cell values were proportions over the total.

To calculate risk, relative counts of oil tracks were, separately, multiplied by relative whale density values and relative probability of occurrence values, in each grid cell. Risk scores in each resulting surface were standardized by divided by the maximum value in the surface, to permit comparison between risk maps. Risk surface and oil track counts surface maps were created in QGIS (QGIS Development team, 2018).

4.3 RESULTS

For longer scenarios, with 60 days for time length, simulations took about 1 hour and 15 minutes to be completed in a computer equipped with a Intel i7 processor (Intel® Core™ i7-4510U CPU @ 2.00–2.60 GHz). Simulations for time lengths of 30 days took about 26 minutes and for the shorter time length (i.e., 15) about 16 minutes.

As expected, longer oil tracks were simulated over longer time periods resulting in wider areas of predicted oil occurrence (compare rows in Figure 4.3 and Figure 4.4 for simulations run for 15, 30 and 60 days). There were some differences in simulated tracks between different months within years, for tracks from fields at around 10ºS of latitude and for those to the south of 20ºS (compare columns in Figure 4.3 and Figure 4.4 for simulations in August and September in 2008 and in 2012). The area south of 20ºS corresponds to the region with most of the fields considered in this study (Figure 4.1). For fields around 10ºS, oil hypothetically spilled in September of 2008 and 2012 was predicted to consistently go to more offshore waters than oil spilled in August, which becomes more evident with increasing time lengths. For fields to the south of 20ºS, distribution of oil tracks is denser offshore in August than September. Despite different figures between months and years, including differences between simulations for August

from 2012 to 2016 (Figure 4.5), oil was consistently predicted over the continental shelf and in offshore regions at the southern limit of the area considered here.

Oil tracks simulated from fields in the northern coast of Brazil (north of 5ºS) consistently predicted that spilled oil would travel westwards along the northern Brazilian continental shelf, within the time frames and years considered here. For fields along the north-eastern coast of Brazil (~10ºS), simulated oil tracks varied considerably; oil was simulated to disperse both north and southwards along the continental shelf, and to nearby offshore waters (Figure 4.3, Figure 4.4 and Figure 4.5). However, due to the small number of fields in that area, oil tracks showed relatively limited spread. Oil tracks simulated from fields located in the south-eastern Brazilian coast, south of 20ºS, were restricted to the southern portion of Abrolhos Bank and further south, within the time considered in the different scenarios, and did not extend to the core inner portions of the Bank (Figure 4.1). Oil tracks were consistently predicted on the portion used by humpback whales in Brazil, over the continental shelf.

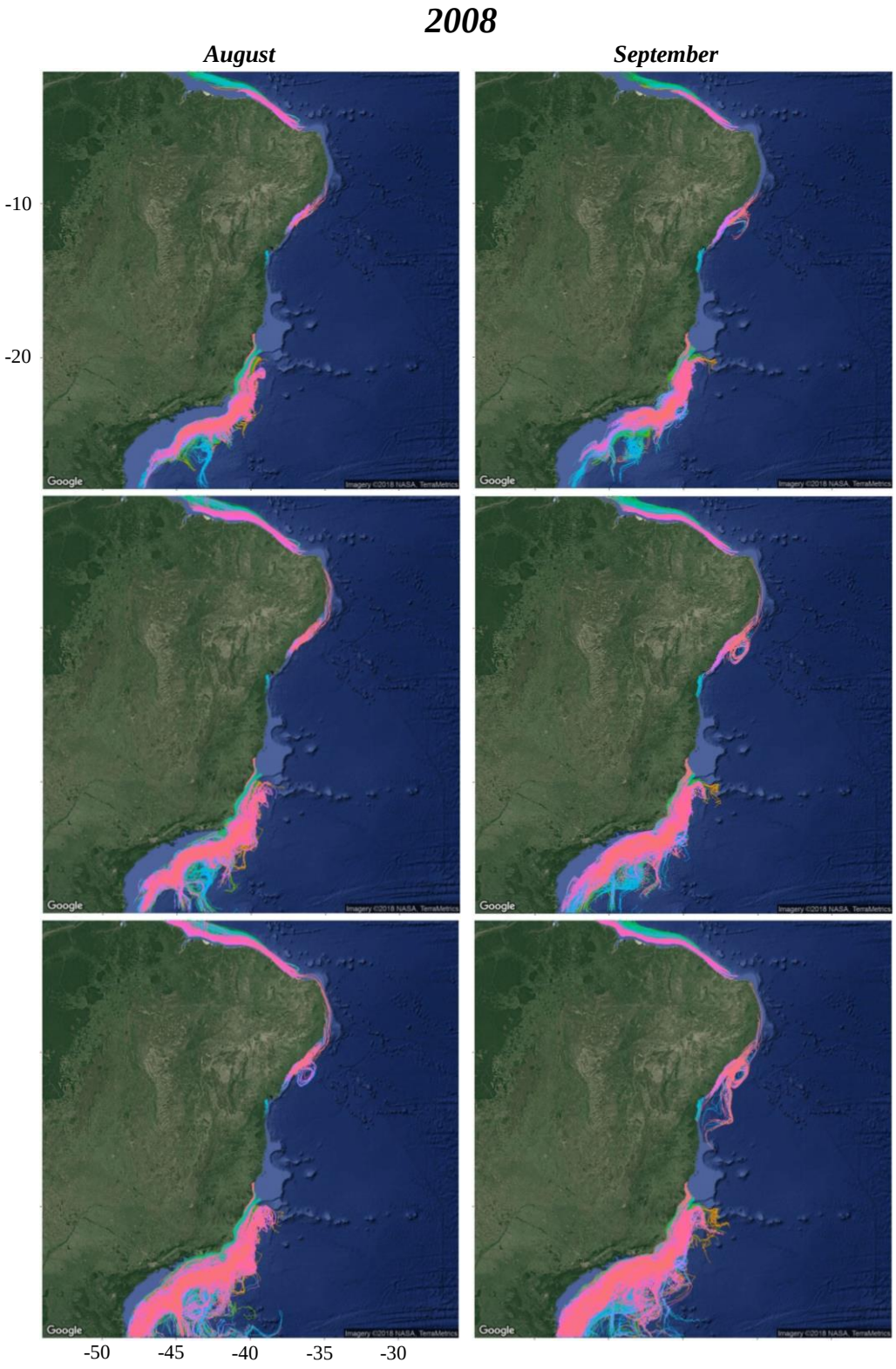


Figure 4.3. Oil tracks simulated for 2008, with spills starting on 1 August and 1 September. Different rows represent different length of dispersion: 15 days long (top), 30 days long (middle) and 60 days long (bottom). Different colours represent different tracks within tracks from the same source.

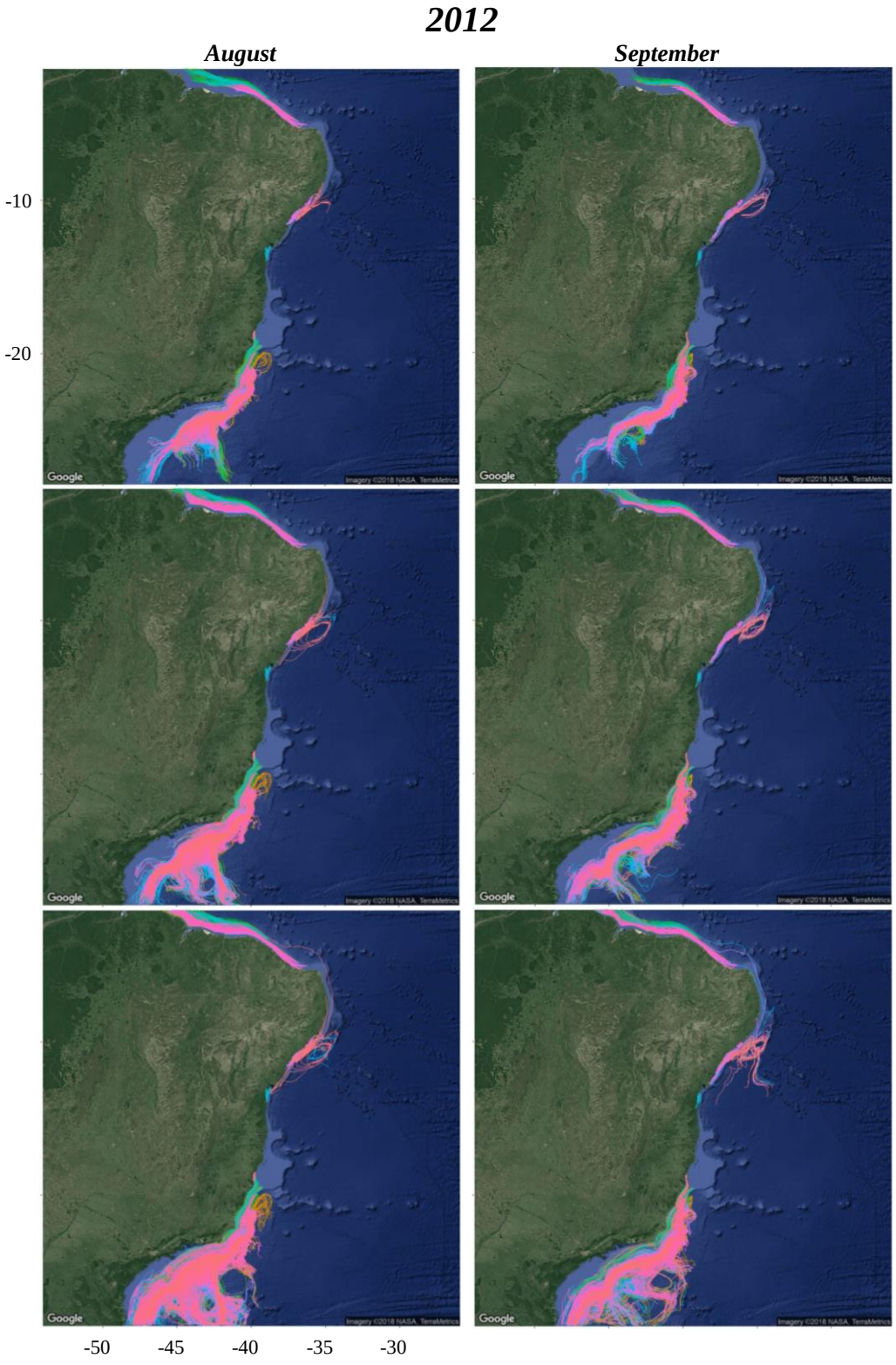


Figure 4.4. Oil tracks simulated for 2012, with spills starting on 1 August and 1 September. Different rows represent different length of dispersion: 15 days long (top), 30 days long (middle) and 60 days long (bottom). Different colours represent different tracks within tracks from the same source.

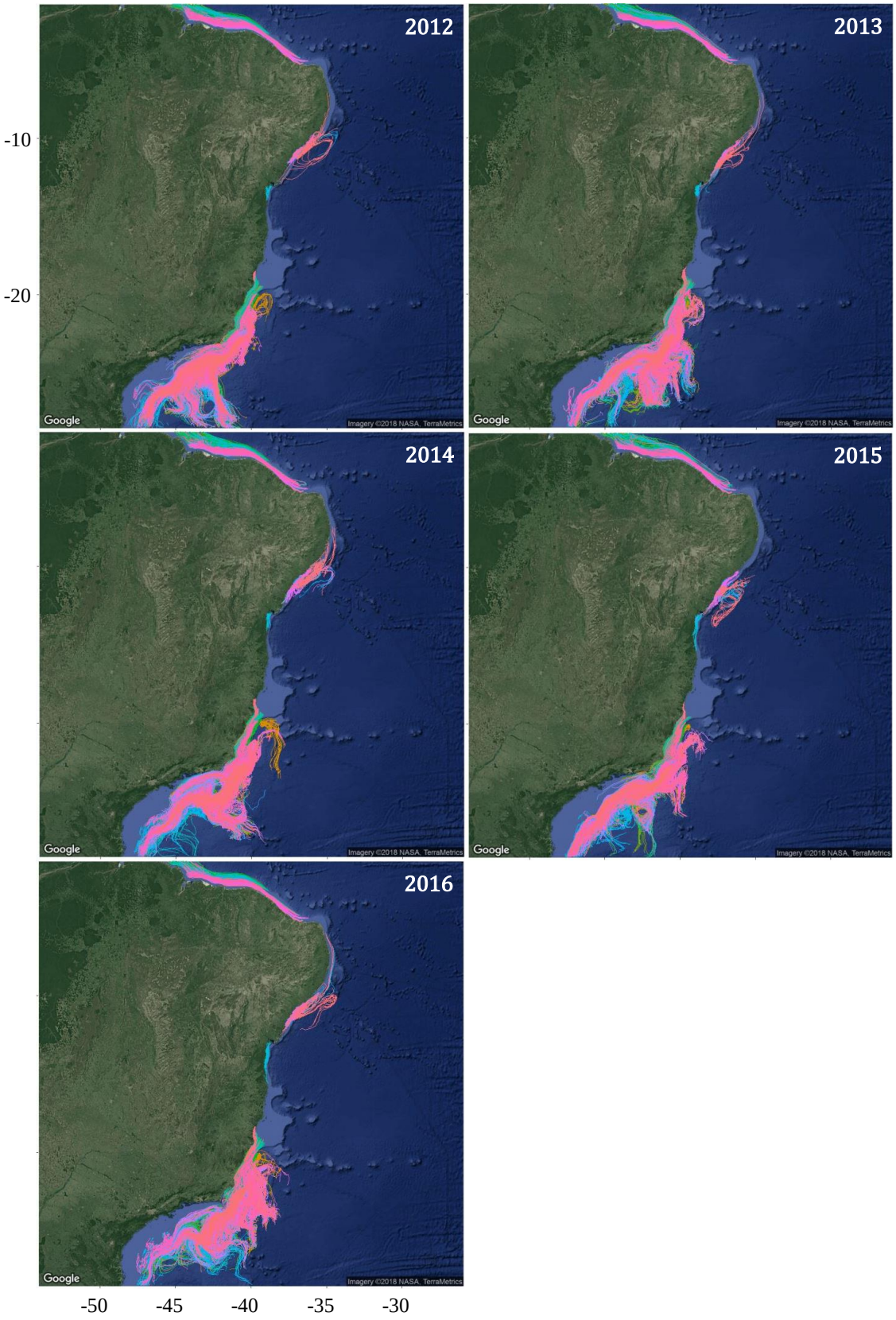


Figure 4.5. Oil tracks simulated for 2012 to 2016, with spills starting on 1 August. Different colours represent different tracks within tracks from the same source.

4.3.1 Risk

For the line transect grid, track counts in the 1,910 grid cells ranged from zero to 436, with most cells not being crossed by any oil track (n = 1,203). For those cells including at least one track crossing (n = 707), about 80% had more than 100 tracks (n = 556) (Figure 4.6). Oil track counts were much higher in the south of the survey area (Figure 4.7). Numbers for the telemetry grid slightly differed in the proportions mentioned above because that grid was slightly smaller and had fewer total cells (n = 1716).

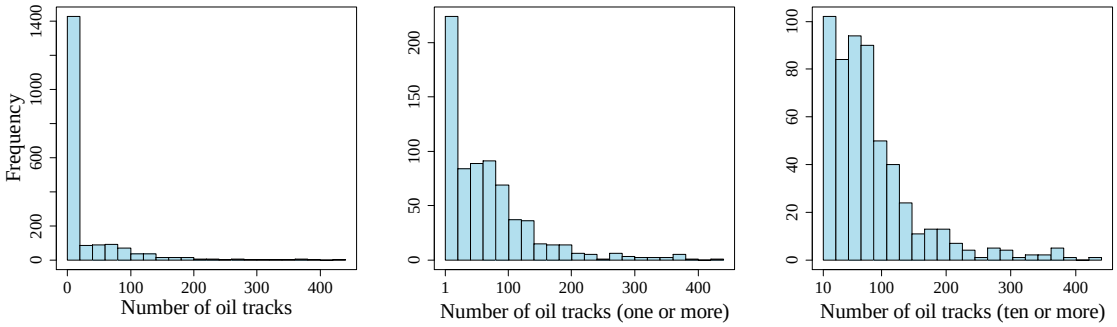


Figure 4.6. Oil track count frequencies considering all grid cells (left), grid cells with at least one track crossing it (middle) and with at least 10 tracks (right), for the line transect grid.

Most of the grid cells in both grids had risk scores of zero (1,203 in the line transect grid and 1,031 in the telemetry grid). High risk scores were calculated for the southern portion of the study area, with highest densities of oil tracks calculated for the extreme south (Figure 4.7 and Figure 4.8).

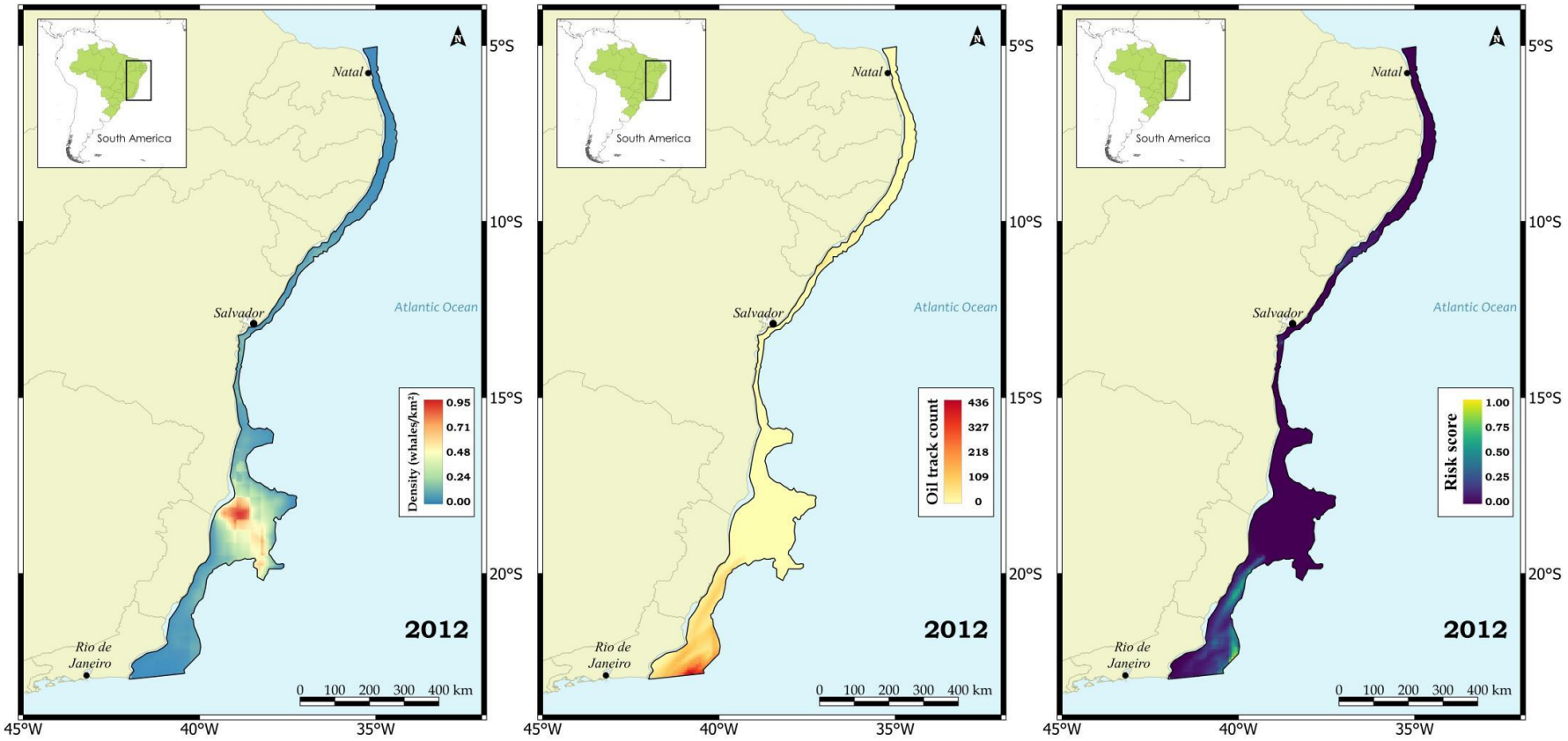


Figure 4.7. Surface maps of whale density (left; Figure 2.10), oil track counts (middle) and relative risk (right) for spills simulated for August 2012.

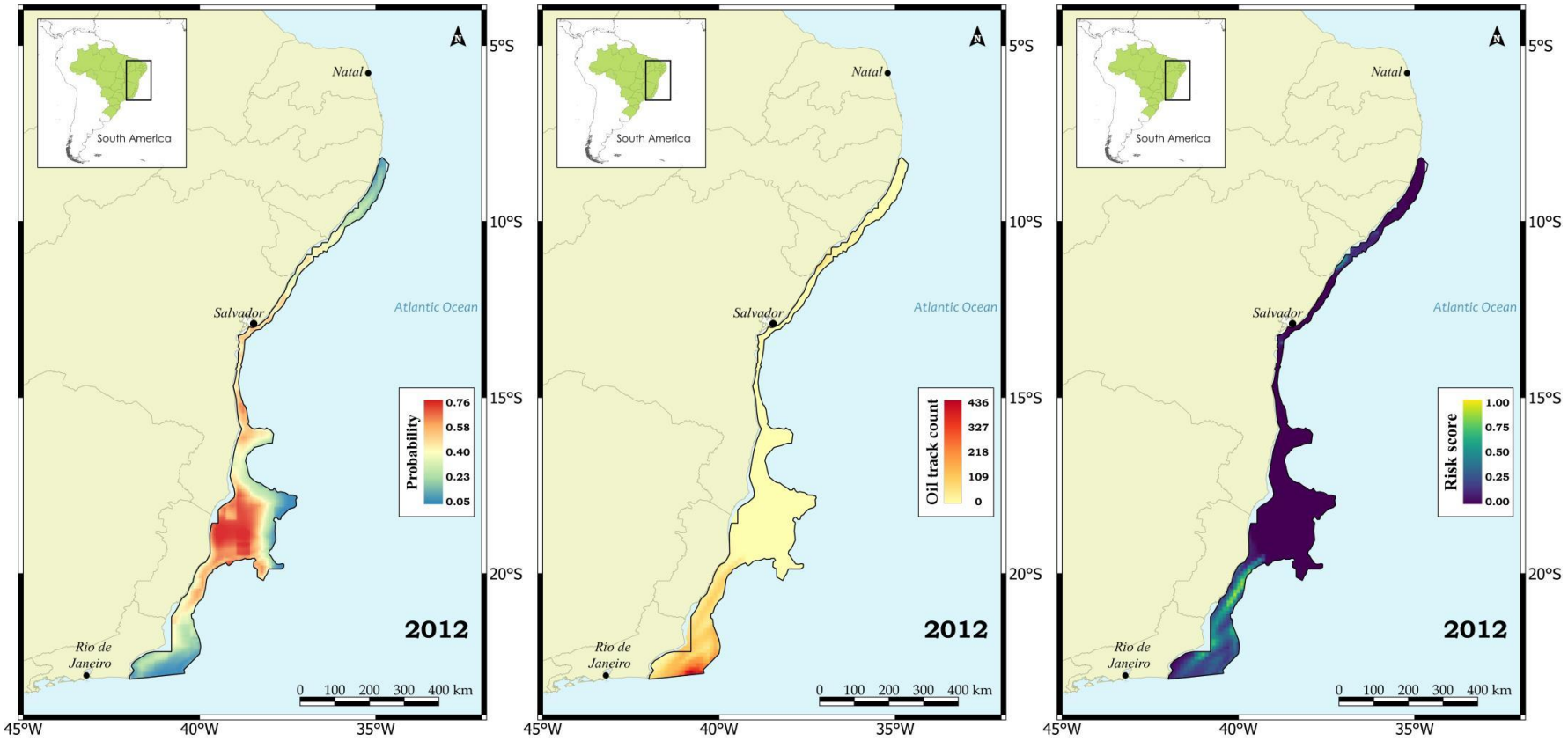


Figure 4.8. Surface maps of probability of occurrence of whales (left; Figure 3.3), oil track counts (middle) and relative risk (right) for spills simulated for August 2012.

4.4 DISCUSSION

The described method provides a fast and easy-to-implement procedure to investigate areas where humpback whales may be at risk of being impacted from oil spills from multiple sources along the Brazilian coast. It has the potential to generate information that may be used to guide management actions for the oil industry in Brazil, to avoid impacts from its activities on the humpback whale population wintering along the country's coast.

The present method relies on strong assumptions about the dispersion of oil. For example, it assumes that oil at the surface disperses under the influence of the same current speed for an entire day. The water motion in any part of the ocean is likely much more complex than that (Manabe, 1969). However, the simplicity adopted here permits the method to be an easy way to calculate possible scenarios of impact to populations for which a distribution surface is available. The simulation of grouped oil tracks to provide the variation in oil dispersion within a month of possible spills (e.g., different simulated oil tracks starting on different days form the same source) was adopted so conclusions should be made at monthly scales. This means that, if the current characteristics have little or no variation during the considered time period (Figure 4.2), that would have produced similar oil dispersion tracks from the same source. Because ocean currents are indubitably more dynamic than considered here, the range of oil tracks could be different if shorter time steps in oil simulation were used. Therefore, the conclusions presented here are best interpreted at wide scales in space and time.

It is important to note that oil does not only travel at the surface of the sea; different densities of the spilled oil may disperse and sink under the influence of other characteristics of the water (Adcroft et al., 2010). For example, oil partially dispersed in the water column could travel under the influence of different currents at different depths. A possible way to evaluate the accuracy of the present method, for dispersion at surface only, would be by using tracked floating devices (Molinari et al., 1990; Cui and Zhao, 2018) and comparing simulated oil tracks from the method to observed dispersion tracks from such devices.

Alternative models of oil dispersion at sea, in which oil dispersion is additionally affected by factors other than surface currents, incorporate different levels of higher complexity (Adcroft et al., 2010; Spaulding, 2017). More complex models often include a

combination of environmental and oil characteristics, such as wind, bathymetric features, type and density of oil, water temperature and other characteristics (Spaulding, 2017). A similarly simple alternative to the present method to approach the risk of impact from oil spills in animals distributed at sea would be to predict where the oil would have spilled from, given that it was hypothetically found in a specific place. This alternative would work using similar calculation steps for the oil tracks used here, however backwards. For example, the path that dispersing oil would have taken if it was hypothetically found where whales are at their highest density, could be calculated. If future current speeds can be predicted for the area of interest, perhaps through monitoring present and past currents, this approach could possibly identify areas where oil sources would present risk to the animals and guide the allocation of oil fields. This approach was not investigated here, but would complement the present study and conclusions.

Information on the overlap in spatial distribution of animal density and human activities has been used to assess risk of impact to whale populations in a number of studies. However, these typically consider fixed human-made structures in the context of existing or planned constructions (Martins et al., 2013), or consistently recurring presence of human activities in the same places, such as large vessels using defined shipping routes (Nichols and Kite-Powell, 2005; Martins et al., 2013). The method presented here partially accommodates the dynamic nature of oil dispersion.

The risk measurement used here was assessed by multiplying relative whale densities and oil track counts across the study area. This caused the large majority of grid cells to have a risk score of zero (Figure 4.6), corresponding to those cells which had no oil tracks crossing, because no cells had zero estimated density of whales.

Regardless of which of the two maps used to represent whale distribution is used, the present method suggests that humpback whales in Brazil are at greater risk of being impacted by oil spills from oil production fields in the south portion of their typical area of occurrence during the breeding season (Figure 4.7 and Figure 4.8). The higher risk for that region calculated here is mainly driven by the high concentration of tracks and not much by whale density (compare maps in Figure 4.7 and Figure 4.8). Determining the consequences of this risk for this growing population is very challenging, because individual level responses to different pressures are not easily extended to the population

scale (Pirotta et al., 2018). In case this becomes a conservation concern in the future, one possible mitigation strategy could be to restrict oil production activities in that region during their breeding season, or during the expected beginning of migration, in the end of winter.

As indicated in Chapter 2, most animals from the humpback whale population in Brazil concentrate in areas far from oil fields during the peak of the breeding season. The present results suggest that the closer whales are to oil fields the higher is the risk of impact from oil spills (Figure 4.1, Figure 4.7 and Figure 4.8), which was expected since oil simulations always started from the same sources across time. While distribution and areas of higher concentration of whales in the wintering grounds suggest that the impact of oil spills is limited during the peak season, higher risk of exposure may occur at the onset of the migration. The southern parts of the typical range of humpback whales in Brazil is used by this population as a pre-migratory habitat prior to initiating their southbound seasonal migration towards the feeding grounds (Zerbini et al., 2006a). Therefore, higher concentration of whales in these habitats later in the season (e.g., November) will lead to a much higher risk to the population should an oil spill occur when whales are about to depart from the wintering grounds. In fact, Zerbini et al. (2006b) suggested that oil exploration activities in *Bacia de Campos* (i.e., high concentration of fields in the south in Figure 4.1) could be an important source of disturbance from the presence of physical structures to whales about to start migration. The present method supports that and adds a quantitative measure of the relative level of potential disturbance from oil spills in that area.

The southern portion of the Abrolhos Bank has been suggested to be the area with highest risk of impact from oil and gas activities to humpback whales, because planned oil exploration activities overlapped with high density of whales there (Martins et al., 2013). The present study indicates that, with the current oil production activities, the highest risk area is further south, with a very small portion at the south of the Abrolhos Bank presenting any risk. Also, less than 8% of the abundance estimated for 2012 corresponded to the area to the south of the Abrolhos Bank (Chapter 2; Bortolotto et al., 2017), further indicating that a small portion of the population may be at risk. However, the present study investigated oil spills from production fields, and there is no disagreement that oil

production activities within the Bank would represent a threat to the whales, because the risk scores calculated in the present study are strongly related to the position of oil sources (compare middle and right maps in Figure 4.7 and Figure 4.8). To help address whether the animals in the Abrolhos Bank could be at high risk of impact from oil spills, the present method could be implemented having the areas for planned oil exploration as hypothetical sources.

Intentionally left blank.

APPENDIX 4.1

Calculating oil dispersion

Production fields downloaded on 5 September 2018
From: http://www.anp.gov.br/exploracao-e-producao-de-oleo-e-gas/dados-tecnicos?view=default

This is part of a PhD chapter of Gui A. Bortolotto
Last update 17 January 2019 (GAB)

```
require("rgdal")
require("plyr")
require("ggplot2")
require("ggmap")
require("raster")
require("dplyr")
require("ddpcr")
setwd("C:/Documents/Project R - Oil/")
```

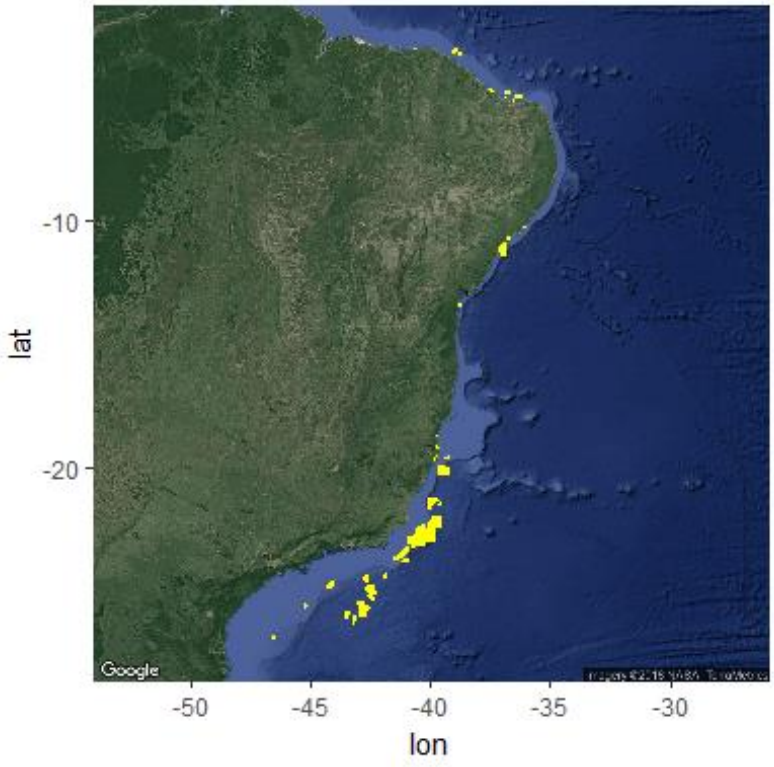
Load necessary components to run the calculation: arrays of current speed values (v and u), time, x and y extents, land maps, production fields.

Current speed data from: Global Ocean Physical Reanalysis product:
GLOBAL_REANALYSIS_PHY_001_030
*Choose the year components to use

```
# Choose year components
# load("DispersionComponents08"); t<- t08; vo.array.utm<- vo08.utm; uo
.array.utm<- uo08.utm
load("DispersionComponents12"); t<- t12; vo.array.utm<- vo12.utm; uo.a
rray.utm<- uo12.utm
load("land_maps")
load("brmap")
fields.sea.sp <- readOGR(dsn=paste0(getwd(),"/gis files- oil"),
                         layer="fields_sea", verbose=F)
#unique(fields.sea.sp@data$id) # 99
fields.sea <- fields.sea.sp@data
```

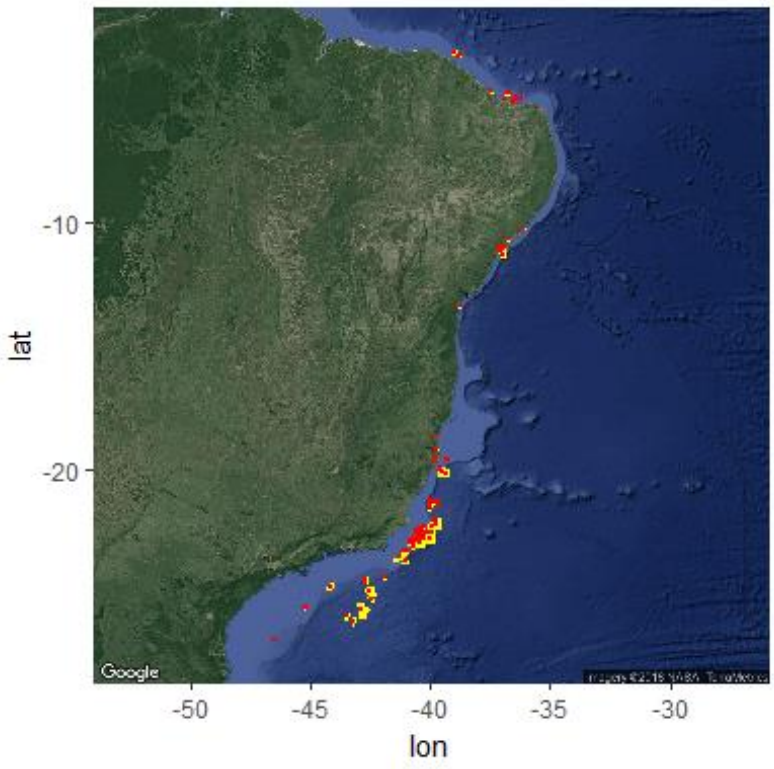
Plot with background map

```
p <- ggmap(br.map)
p <- p + geom_polygon(aes(long,lat,group=id), fill="yellow", col="yell
ow",
                      data=fields.sea)
p
```



Calculate centroids and add to the previous plot

```
centroids <-fields.sea %>% group_by(id) %>% summarise(long=mean(long),
lat=mean(lat))
centroids <- data.frame(centroids)
p2 <- p + geom_point(data=centroids, aes(x=centroids$long, y=centroids
$lat), colour="red", size=0.3, show.legend = F)
p2
```



Dealing with centroids

```
# Project to UTM
centroids.sp <- SpatialPointsDataFrame(cords = cbind(centroids$long,
      centroids$lat), data=centroids, proj4string = crs(land.map))
centroids.sp.utm <- spTransform(centroids.sp, crs(land.map.utm))
# Add x and y to centroids data frame
centroids$x <- centroids.sp.utm@coords[,1]
centroids$y <- centroids.sp.utm@coords[,2]
# Exclude centroids inland
over.centroids <- over(centroids.sp.utm, land.map.utm)
centroids <- centroids[(is.na(over.centroids)),]
# Create a list of locations for centroids
initial.time <-t[1] #format="%d-%m-%Y", tz = "UTC"
```

Choose time span

Example: to evaluate an entire month of possible oil dispersion tracks, use 'time.spam = 30'

```
time.span = 30
# time
pts.time <- as.POSIXct(initial.time, format="%d-%m-%Y", tz = "UTC")
pts.time <- seq(from=pts.time, to=pts.time+ (time.span-1)*24*60*60, by
=24*60*60)
pts<-rep(list(0), nrow(centroids))
# str(centroids)
for(i in 1:nrow(centroids)){
```

```
  temp.id = centroids$id[centroids$id==unique(centroids$id)[i]]
  temp.long = centroids$long[centroids$id==unique(centroids$id)[i]]
  temp.lat = centroids$lat[centroids$id==unique(centroids$id)[i]]
  temp.x = centroids$x[centroids$id==unique(centroids$id)[i]]
  temp.y = centroids$y[centroids$id==unique(centroids$id)[i]]
  pts[[i]] <- data.frame(id=temp.id, long=temp.long, lat=temp.lat,
                         time=pts.time, x=temp.x, y=temp.y)
}
# create a list for locations
df.loc <- data.frame(x=NA,y=NA,time=as.Date("1111-11-11"), id=NA, inde
x=NA)
list.loc <- rep(list(df.loc), nrow(centroids))
```

Calculate oil dispersion

The oil dispersion is calculated considering that oil is carried by the sea surface current, for one day on each step, starting from its source. This means that the oil will disperse in the same direction for an entire day under the influence of the current at the sea surface at that point. On the next step (i.e., next day) it will disperse under the influence of the current at that point and so on. The sources of oil here are the centroids of oil field at the production stage.

Choose time of dispersion

```
time.disperse <- 30 # amount of days of dispersion for each track
```

Dispersion calculation

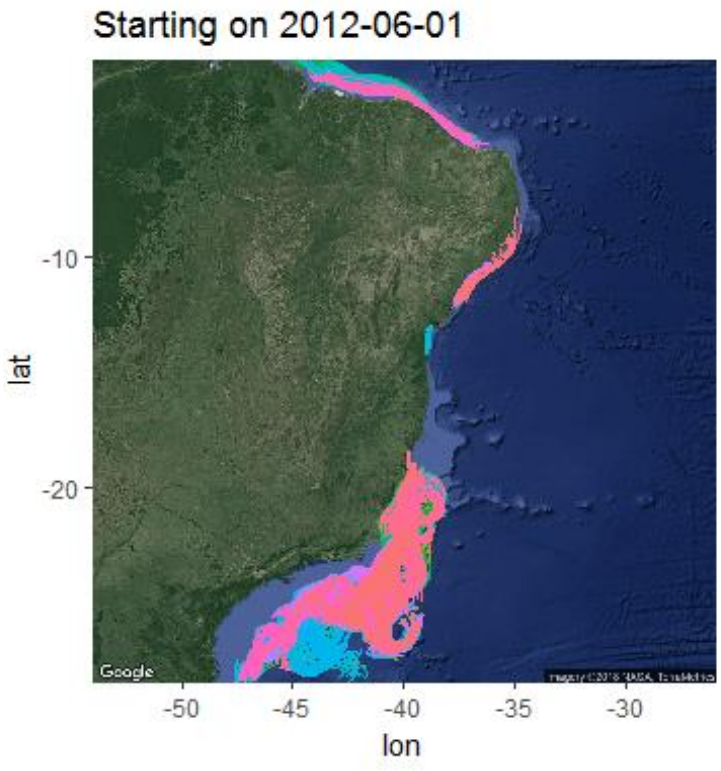
```
# # # RUN
StTime<-proc.time() # to calculate the time it takes to run
for(k in 1:nrow(centroids)){
  # for k iterations, feed 'locations' and

  # 'pt' with elements in 'list.loc' and 'pts'
  locations = list.loc[[k]]
  pt = pts[[k]]
for(i in 1:nrow(pt)){
  init.time <- pt$time[i]
  init.loc <- pt[i,c("x","y")]
  init.data<-data.frame(x = init.loc$x, y = init.loc$y,time=init.time,
        id=paste0(unique(centroids$id)[k],"a",i), index=1)
  locations<-rbind(locations, init.data)
  rvo.utm<-raster(vo.array.utm[,,which(t==init.time)],
        xmn=min(extent.x), xmx=max(extent.x),
        ymn=min(extent.y), ymx=max(extent.y),
        crs="+proj=utm +zone=24 +south +ellps=aust_SA +units=m +no_def
s")
  ruo.utm<-raster(uo.array.utm[,,which(t==init.time)],
        xmn=min(extent.x), xmx=max(extent.x),
        ymn=min(extent.y), ymx=max(extent.y),
        crs="+proj=utm +zone=24 +south +ellps=aust_SA +units=m +no_def
s")
```

```
  e.v<- extract(rvo.utm,init.loc)
  e.u<- extract(ruo.utm,init.loc)
  for(j in 1:(time.disperse)){
    if(j==1){ # the 1st new location (i.e., the 2nd of the track)
      # here the new step (1 day = 24*60*60 sec) is calculated
      new.loc<-init.loc+(c(e.u,e.v)*86400)
      new.time<-init.time + 86400
      new.data<-data.frame(x=new.loc$x,y=new.loc$y,time=new.time,
        id=paste0(unique(centroids$id)[k],"a",i), index=i+1)
      locations<-rbind(locations, new.data)
     }else{
      new.time<-new.data$time + 86400
      rvo.utm<-raster(vo.array.utm[,,which(t==new.time)],
        xmn=min(extent.x), xmx=max(extent.x),
        ymn=min(extent.y), ymx=max(extent.y),
     crs="+proj=utm +zone=24 +south +ellps=aust_SA +units=m +no_defs")
      ruo.utm<-raster(uo.array.utm[,,which(t==new.time)],
        xmn=min(extent.x), xmx=max(extent.x),
        ymn=min(extent.y), ymx=max(extent.y),
     crs="+proj=utm +zone=24 +south +ellps=aust_SA +units=m +no_defs")
      e.v<- extract(rvo.utm,new.loc)
      e.u<- extract(ruo.utm,new.loc)
      if(is.na(e.v) | is.na(e.u)) {
        break
      } else {
      new.loc <- new.data[1:2]+(c(e.u,e.v)*86400)
      new.data<-data.frame(x =new.loc$x, y = new.loc$y, time=new.time,
        id=paste0(unique(centroids$id)[k],"a",i), index=j+1)
      locations <- rbind(locations, new.data)
      }
      }
     }
   }
   list.loc[[k]]<-locations[-1,]
}
#close(pb)
EnTime<-proc.time()
c(EnTime-StTime)['elapsed'] # show time it took to run (secs)

## elapsed
## 1133.48

# # # END RUN
```

Plot result

```
p <- ggmap(br.map)
p3 <- p
for(i in 1:length(list.loc)) {
  the.locations <- data.frame(list.loc[i])
  loc.lon.lat <- SpatialPointsDataFrame(the.locations[,1:2], data=the.
locations[,3:5],
```

```
          proj4string = CRS( "+proj=utm +zone=24 +south +ellps=aust_SA
+units=m +no_defs"))
  loc.lon.lat <-spTransform(loc.lon.lat, "+proj=longlat +datum=WGS84")
  loc.lon.lat.points <- data.frame(lon=loc.lon.lat@coords[,1],
          lat=loc.lon.lat@coords[,2], id=loc.lon.lat@data$id)
      p3 <- p3 + geom_path(data=loc.lon.lat.points, aes(x=lon, y=lat,
          group=id, colour=factor(id)),show.legend = F)
      }
      p3 <- p3 + ggtitle(paste("Starting on",t[1]))
      p3
```



Unlist and save locations

```
dispersion.locations <- do.call("rbind", list.loc)
write.csv(dispersion.locations, file="dispersion_locations.csv")
#save(dispersion.locations,list.loc, file=oil_output)
```

CHAPTER 5

REVISITING AND UPDATING THE ASSESSMENT OF SOUTHWESTERN ATLANTIC HUMPBACK WHALE POPULATION STATUS

ABSTRACT

Humpback whales from the population breeding in Brazilian waters were exploited by commercial whaling almost to the point of extinction in the mid-nineteenth century. The conservation status of this population was assessed in the past, however through relatively complicated methods. An updated assessment, including recently available data, is also currently lacking. In this chapter, a previously published Bayesian population dynamics model, describing the population's trajectory from 1900 and projecting to 2040 (Zerbini et al., 2011), was implemented in a state-space model framework and using a different, more standard fitting method (i.e., Markov chain Monte Carlo [MCMC], as opposed to Sampling Importance Resampling) to: 1) replicate the "base case scenario" presented in that study and 2) update the model with different and additional new data. New data available for updating the model were abundance estimates for 2008 and 2012, from Chapter 2, and new demographic information. Objectives here were to investigate if the model could be replicated using the different fitting method and model framework and the effect of using different data on model outputs. Quantities of interest included: pre-exploitation population carrying capacity (K), maximum population growth rate (r_{max}), maximum level of depletion ($Nmin/K$) and current level of depletion (N_{2019}/K). Zerbini's base case was successfully replicated using MCMC. The updated model provided more precise estimates for population sizes along the period considered (1901-2040) and indicated that carrying capacity and maximum level of depletion were lower than estimated in Zerbini's base case. However, the posterior 95% credible intervals of parameters in the new model overlap with those of the previous study. The results suggest that the population could reach carrying capacity sooner than previously expected, i.e., within the next two years.

5.1 INTRODUCTION

Whaling during the twentieth century may be considered the greatest human exploitation of wildlife in history (Clapham et al., 2008). Modern whaling operations reduced most humpback whale populations in the Southern Hemisphere to small fractions of their size before commercial exploitation (Findlay, 2000), including the population breeding in Brazilian waters. Despite the International Whaling Commission protection regulations established in 1966, soviet catches in the area until the 1970s (Zemsky et al., 1996) further decreased population numbers. Investigating the consequences and level of impact of such activities on whale populations is important to understand their present conservation status and to evaluate the need for and to guide management actions.

Humpback whales in the southwestern Atlantic (WSA) spend winter and spring months in the coastal waters of Brazil, where pregnant females give birth and sexually mature animals mate (Martins et al., 2001; Andriolo et al., 2010). These animals migrate every year between their breeding grounds in Brazil and their feeding grounds around South Georgia and South Sandwich Islands (Zerbini et al., 2006; Engel and Martin, 2009). It is well documented that the population size has increased considerably in recent years (Andriolo et al., 2010; Zerbini et al., 2011; Bortolotto et al., 2016a; 2017; Pavanato et al., 2017; Wedekin et al., 2017). Wedekin et al. (2017) estimated that this population has grown extremely fast for cetaceans with similar life histories, having increased 12% per year from 2002 to 2011. The increase in whale numbers coincides with an increase in human activities in their area of distribution during their breeding season (Chapter 4). It has been estimated that around 20,000 humpback whales were present in the Brazilian breeding ground in 2012 (Chapter 2; Bortolotto et al., 2017). This figure represents 80% of what was estimated in Zerbini et al. (2011) as the population size before 1900, when commercial whaling in the area started using modern methods, such as explosive harpoons and steam-powered boats (Clapham and Baker, 2018). Zerbini et al. (2011) developed a model of the population trajectory projected back in time, fitted to an abundance estimate from an aerial survey (Andriolo et al., 2010) in 2005. However, that aerial survey estimate is likely biased low (Bortolotto et al., 2016a) and the assessment of this population's conservation status can benefit from new data currently available. In earlier efforts to assess this population's conservation status (e.g. Zerbini et al., 2005), the International Whaling Commission Scientific Committee (IWC-SC) had proposed

different scenarios for numbers of catches in their feeding grounds, to account for uncertainty about where those areas could be (IWC, 1998). Information from satellite tracking and animal identification (Engel and Martin, 2009; Stevick et al., 2006; Zerbini et al., 2006; 2011) support the "Core hypothesis" (IWC, 2011), that waters near South Georgia and South Sandwich Islands represent their main feeding grounds.

For quantitatively investigating biological parameters that are observed with error, such as when abundance estimates are used in population dynamics models, statistical tools that account for such errors are important to provide robust outputs. Bayesian state-space models applied to population dynamics (Gimenez et al., 2009; Newman et al., 2014) allow two distinct components of a model to be specified: the process component, which describes the hypothesized, but unobserved, population dynamics through the relation between biological parameters of interest; and the observation component, which can be used to describe uncertainty in the data observed (Gimenez et al., 2009). In their study on WSA humpback whale conservation status assessment, Zerbini et al. (2011) used Sampling-Importance-Resampling (SIR; e.g., McAllister et al., 1994) to estimate parameters in a density-dependent, sex and age-aggregated population dynamics model, such as that described by Pella and Tomlinson (1969). A backwards approach was used (Butterworth and Punt, 1995), starting from a recent abundance estimate (in this case for 2005) to project back in time and estimate the population size in 1901, which is considered to represent the pre-exploitation carrying capacity of the population (K). However, small scale whaling in the area before that year (Morais et al., 2017), may have previously depleted the population.

Because Zerbini et al. (2011) adopted a deterministic underlying population model (equivalent to the process component in state-space models), strictly it should not be considered a state-space model because in that class of models both the process and observation components are stochastic. Here, an alternative Bayesian method to fit that model in a state-space framework was used to allow for a stochastic process component: Markov chain Monte Carlo (MCMC) simulations using the Metropolis algorithm (Metropolis et al., 1953; Geyer, 1992; Kruschke, 2014).

In this chapter the assessment of the conservation status of humpback whales from the WSA population was revisited by implementing and updating a previously published

population dynamics model (Zerbini et al., 2011), using a different method for model fitting. The objectives were 1) to reproduce the results described in Zerbini et al. (2011); 2) to investigate model outputs resulting from inclusion of new population abundance and growth information; and 3) investigate model structure, including the effect of prior distributions (based on expectations *a priori* about parameter values) and data on the outputs. Results presented here can contribute to the understanding of the impact that modern whaling had on this population and can assist informing the need for management actions.

5.2 MATERIAL AND METHODS

A Bayesian state-space model, formed by a process component (i.e., population model) and an observation component (i.e., observation model) was applied to investigate humpback whale population dynamics. The population dynamics model presented here was based on the "base case scenario" described in Zerbini et al. (2011; Figure 5.1), which assumes that the WSA humpback whale population was at carrying capacity at the beginning of the 20th century, before modern whaling activities. Zerbini et al. (2011) used a Sampling-Importance-Resampling (SIR) algorithm for inference (McAllister et al., 1994); here a Markov chain Monte Carlo algorithm (Geyer, 1992), specifically a Metropolis algorithm (Metropolis et al., 1953; Kruschke, 2014) was used. Custom functions were created in software R (version 3.4.2; R Core Team, 2017) to implement two models: 1) the "base case model", to replicate the base case scenario from Zerbini et al. (2011); and 2) the "updated model", in which different and new data were used to update the base case model. R code for implementing the models is provided in Appendices 5.1 and 5.2.

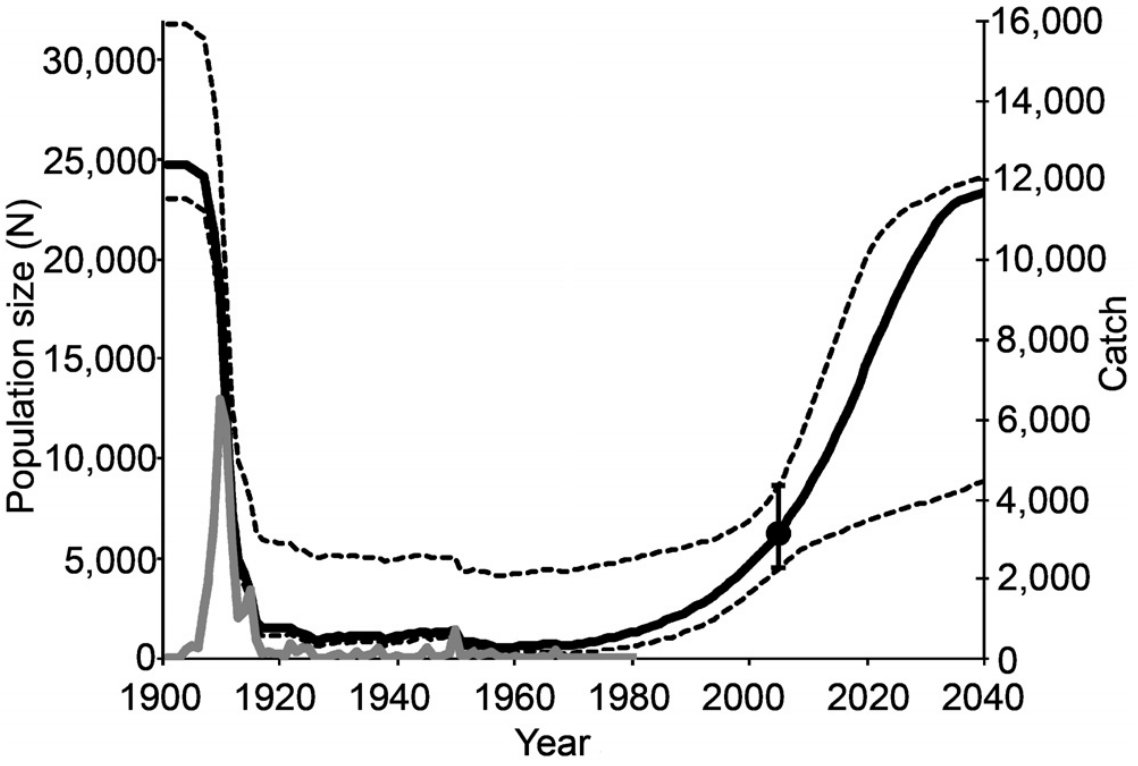


Figure 5.1. Extracted from Zerbini et al. (2011). Population trajectory fitted for Zerbini's base case scenario. Black lines represent the posterior median (solid line) and 95% probability intervals (dashed lines) estimated for population size (left-hand side axis). Observed abundance in 2005 is indicated by a black circle with error bars (95% confidence interval). The grey line represents catches (right-hand side axis) (Table 5.2).

5.2.1 Population model

The population dynamics model (Zerbini et al., 2011) is a density-dependent, sex and age-aggregated generalised logistic equation (Pella and Tomlinson, 1969) that describes the population trajectory from the year 1901, projecting to 2040. The deterministic structure of the model is

$$N_{t+1} = N_t + N_t \cdot r_{max} \cdot \left[1 - \left({}^{N_t}/_{K}\right)^{z}\right] - C_t$$

where N_t is the population size at year t; r_{max} is the maximum net recruitment rate; K is the carrying capacity; z is a shape parameter fixed at 2.39 (Butterworth and Best, 2004), that determines the relative population level at which maximum net recruitment occurs; C_t is catch at year t. In this model, the population size at the beginning of the modelled period is assumed to be equal to the carrying capacity K, which is assumed to be constant over time.

Quantities of interest derived from the fitted models included the population size at the beginning of the period modelled (defined as the carrying capacity, *K*), the maximum depletion level (*min[N]/K*), which represents the lowest population level reached, and the current depletion level of the population ($N_{current}/K$). Depletion levels were calculated by dividing estimated values of *N* by estimated values of *K*. The year in which the minimum population size (*min[N]*) occurred was that which had the smallest median value of *N*.

5.2.2 Data

In the base case model, the same data considered in Zerbini's base case scenario were used: abundance estimate for 2005, estimated growth rate for 1995-1998 and the "core hypothesis" series of catches (Table 5.1). In the updated model, population abundance estimates for 2008 and 2012 from Chapter 2, which were derived from spatial models applied to line transect data collected in ship-based surveys, were also included. New information on population growth rate (Wedekin et al., 2017) was also used in the updated model (Table 5.1). The abundance estimate derived from the aerial line transect survey data for 2005 (Andriolo et al., 2006) considered by Zerbini et al. (2011) as observed population size, was not used in the updated model because that estimate is not consistent with those from Chapter 2 and is believed to be biased low (see section 5.4). Because abundance and population growth rate estimates have associated error, an observation component for the present state-space model was specified (see section 5.2.3). Catch data were assumed to be known, but see Discussion (section 5.4) for alternatives.

Table 5.1. Data used for modelling population dynamics in the base case model and the updated mode. The "core hypothesis" catch series is given in Table 5.2.

Information	Data	Source	Base case	Updated
Abundance 2008	14,264 (CV = 0.084)	Chapter 2		*
Abundance 2012	20,389 (CV = 0.071)	Chapter 2		*
Abundance 2005	6,251 (CV = 0.17)	Andriolo et al. (2006)	*	
Growth rate (*r*) for 1995-1998	0.074 (CV = 0.446)	Ward et al. (2011)	*	*
Growth rate (*r*) for 2002-2011	0.1135 (CV = 0.115)	Wedekin et al. (2017)		*
Catch series	"Core hypothesis"	Zerbini et al. (2011)	*	*

The catch series used here represents numbers of whales killed during modern whaling activities, defined as the "core hypothesis" (IWC, 2011) presented in Zerbini et al. (2011; Figure 5.2, Table 5.2). The core hypothesis considers that humpback whales killed at feeding grounds between longitude 70-20ºW and latitude 40-50ºS , plus those killed between 50-20ºW and to the south of 50ºS, excluding Falkland catches, belonged to the population breeding along the Brazilian coast (Figure 5.2). Whaling records from catcher boats operating from whaling stations in Brazil (Williamson, 1975), and records of catches made by a Soviet pelagic fleet in the Abrolhos Bank and in offshore areas along the central coast of South America, including illegal catches (Zemsky et al., 1996; Allison, 2006), represent breeding ground catches. The catch series used here was reviewed in Zerbini et al. (2011) and is reproduced in Table 5.2.

Table 5.2. Core hypothesis catches reproduced from Zerbini et al. (2011; Table 4).

Year	**Catches**	**Year**	**Catches**	**Year**	**Catches**	**Year**	**Catches**
1904	180	**1920**	102	**1936**	105	**1952**	34
1905	288	**1921**	9	**1937**	242	**1953**	140
1906	240	**1922**	364	**1938**	0	**1954**	44
1907	1261	**1923**	133	**1939**	2	**1955**	96
1908	1849	**1924**	266	**1940**	36	**1956**	167
1909	3391	**1925**	254	**1941**	13	**1957**	61
1910	6468	**1926**	7	**1942**	0	**1958**	16
1911	5832	**1927**	0	**1943**	4	**1959**	15
1912	2881	**1928**	19	**1944**	60	**1960**	27
1913	999	**1929**	51	**1945**	238	**1961**	13
1914	1155	**1930**	107	**1946**	30	**1962**	24
1915	1697	**1931**	18	**1947**	35	**1963**	12
1916	447	**1932**	23	**1948**	48	**1964**	0
1917	121	**1933**	132	**1949**	83	**1965**	52
1918	129	**1934**	57	**1950**	698	**1966**	0
1919	111	**1935**	48	**1951**	45	**1967**	189

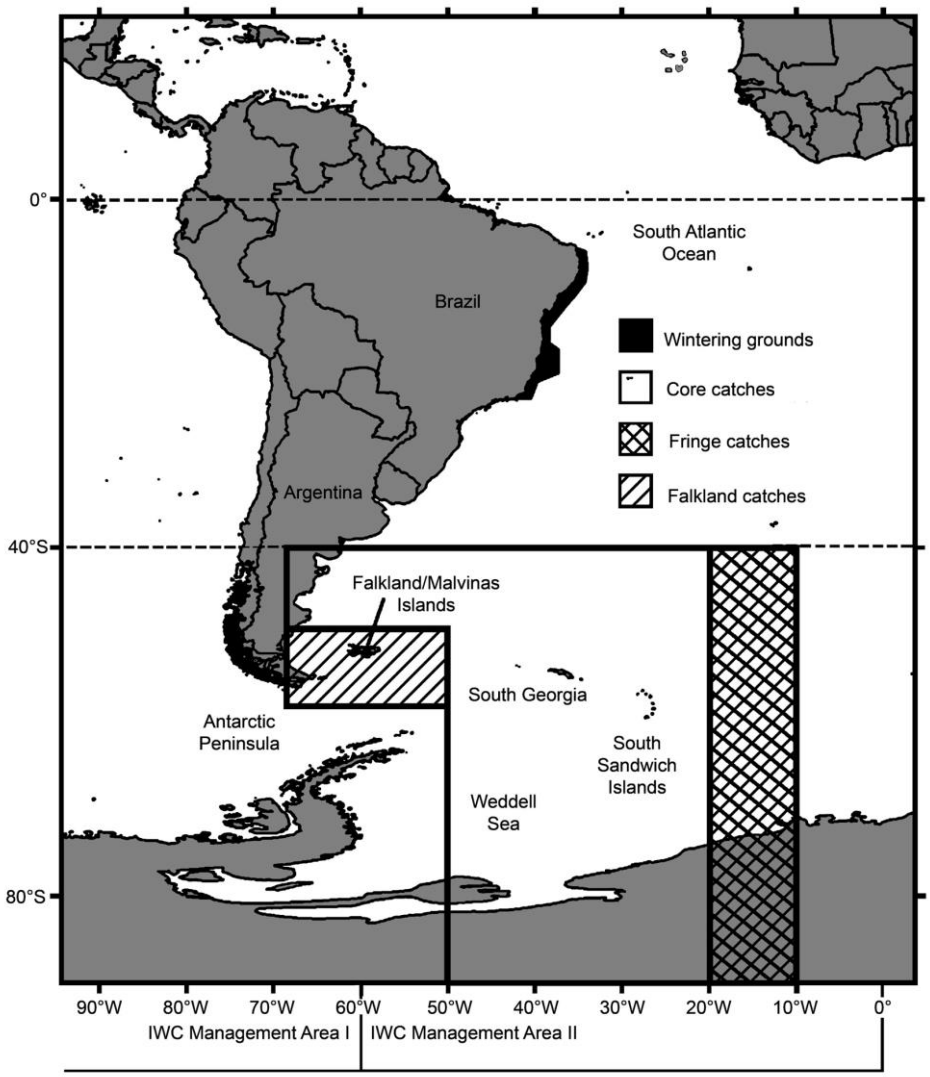


Figure 5.2. Figure extracted from Zerbini et al. (2011) showing limits for different catch allocation hypotheses/scenarios investigated in that study for feeding grounds i.e., to the south of 40ºS. The only feeding ground catch allocation scenario considered in the present study is the "Core catches".

To investigate the sources of differences from the base case model to the updated model, five additional models ("Model checks") were run:

- Model check 1: the base case model was updated only with the additional information on population growth rate from Wedekin et al. (2017);
- Model check 2: the base case model was updated by replacing the abundance estimate for 2005 (Andriolo et al., 2006) by the abundance estimate for 2008 (Chapter 2);
- Model check 3: the base case was updated by replacing the abundance estimate for 2005 by the abundance estimate for 2012 (Chapter 2);

- Model check 4: the base case was updated by replacing the abundance estimate for 2005 by abundance estimates for both 2008 and 2012 (Chapter 2);
- Model check 5: the base case was updated by not including any abundance or population growth rate data. This enables discovery of the effect of the prior distribution specifications and model structure on the outputs of interest.

5.2.3 Observation model

In the observation component of the state-space model it was assumed that the observation on population growth rate ($\hat{r}$) between two years (*x* and *y*) followed a normal distribution:

$$\hat{r}_{x,y} \sim Normal(r_{x,y}, \sigma^2_{\hat{r}_{x,y}})$$

where

$$r_{x,y} = \frac{\ln(N_y) - \ln(N_x)}{y - x}$$

and $\sigma^2_{\hat{r}_{x,y}}$ is a variance parameter, assumed known (see growth rate CVs in Table 5.1).

A lognormal distribution was assumed for observed abundance in year t:

$$\widehat{N}_t \sim Lognormal(\ln(N_t), \sigma^2{}_{log\widehat{N}_t})$$

where $\sigma^2{}_{log\widehat{N}_t}$ is another variance parameter that is assumed known (Table 5.1).

The values for annual population size estimated by the population model formed the population trajectory. The log-likelihood (ln(*L*)) of the population trajectory was calculated by summing the negative log-likelihoods of abundance and population growth rate data, where these log-likelihoods are given below.

The log-likelihood of population growth rate *r* was calculated as

$$\ln\left(L_{r_{x,y}}\right) \propto -\ln\left(\sigma_{\hat{r}_{x,y}}\right) + 0.5 \cdot \left(\frac{r^{pred}_{x,y} - \hat{r}_{x,y}}{\sigma_{\hat{r}_{x,y}}}\right)^2$$

where

σ is the standard error;

$\hat{r}_{x,y}$ is the growth rate observed between years *x* and *y*;

$r_{x,y}^{pred}$ is the growth rate predicted by the model between years *x* and *y*;

And constant terms are omitted (hence the proportional sign, $\propto$).

The log-likelihood of the abundance estimate in year (*N*) was calculated as

$$\ln(L_N) \propto -\ln\left(\sigma_{log\widehat{N}}\right) + log\widehat{N} + 0.5 \cdot \left(\frac{logN^{pred} - log\widehat{N}}{\sigma_{log\widehat{N}}}\right)^2$$

where

σ is the standard error;

$log\widehat{N}$ is the natural logarithm of the observed abundance;

$logN^{pred}$ is the natural logarithm of the abundance predicted by the model;

again, constant terms are omitted from the equation

5.2.4 Prior distributions

The same prior distributions for the model parameters as used in the base case scenario of Zerbini et al. (2011; thin lines in Figure 5.3) were also used here: a uniform distribution for the prior on population size in 2005 (N_{2005}), with lower bound 500 and upper bound 22,000; a uniform distribution for the prior on r_{max} with lower bound zero, because negative values for maximum recruitment rate are biologically unreasonable, and an upper bound 0.106, which corresponds to the maximum reasonable value for growth rate for the species (Clapham et al., 2006).

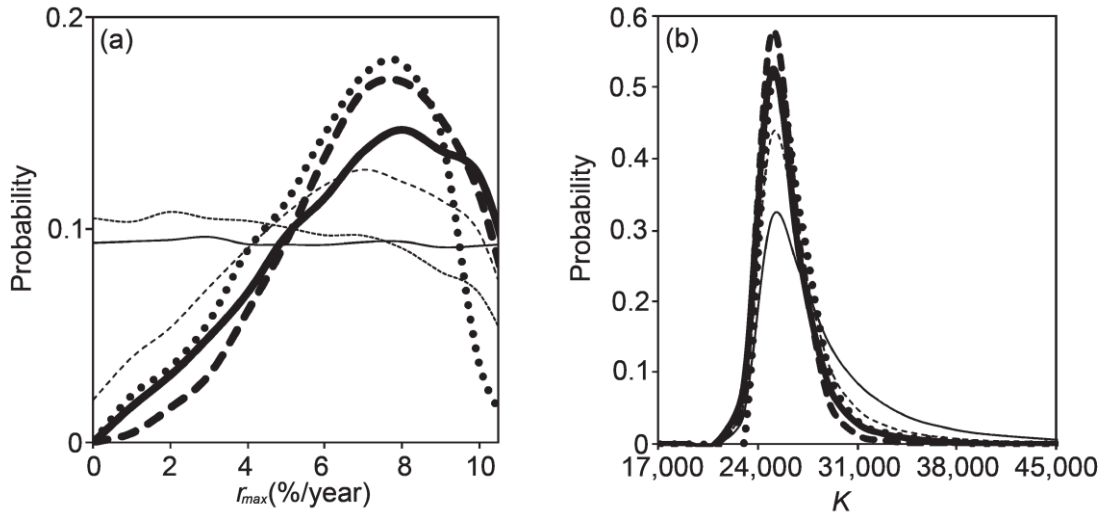


Figure 5.3. Extracted from Zerbini et al. (2011). Posterior (thick lines) and prior (thin lines) probability distributions of r_{max} and K (solid lines represent the base case scenario; other lines represent other scenarios considered in that study). Note the very informative prior on K (thin solid line in the right-hand plot).

The population model and prior distributions on N_{2005} and r_{max} together implied a prior distribution for population size in the first year (N_{1901}), i.e., the carrying capacity, K. A Monte Carlo sampling approach was used to derive this prior: a set of 10,000 independent samples were simulated from the prior distributions on N_{2005} and r_{max}. For each sample, a univariate optimization was conducted using a bisection algorithm to find the value of N_{1901} that, when projected forward using the population model with the known catches and given value of r_{max}, produced the given value of N_{2005}. This approach was also used by Zerbini et al. (2011) as part of their fitting algorithm, and is referred to by them as "backward projection".

5.2.5 Posterior distributions

The joint posterior density of r_{max} and N_{2005} is proportional to the prior multiplied by the likelihood. In this case, with uniform prior distributions on r_{max} and N_{2005}, the posterior density was proportional to the likelihood within the range of the prior distributions on r_{max} and N_{2005}, and zero outside this range. For example, in the base case, using $\pi(r_{max}, N_{2005})$ to denote the posterior distribution for r_{max} and N_{2005}, and working on the log scale

$$\ln\left(\pi(r_{max}, N_{2005})\right) \propto \begin{cases} \ln\left(L_{N_{2005}}\right) + \ln\left(L_{r_{1995,1998}}\right) & , \text{ when } \begin{matrix} 0 \leq r_{max} \leq 0.106 \ \& \\ 500 \leq N_{2005} \leq 22{,}000 \end{matrix} \\ 0 & , \text{ otherwise} \end{cases}$$

5.2.6 Computation

MCMC was used to produce samples from the posterior distribution for the unknown parameters r_{max} and N_{2005}; values for the other quantities of interest (e.g., population size in other years and carrying capacity *K*) were obtained as part of the MCMC computation. To create the Markov chains, a simple Metropolis algorithm (Metropolis et al., 1953) was implemented, employing a bivariate normal proposal centred on the current value of the unknown parameters. A bivariate proposal was used to allow flexibility to incorporate correlation between the posterior distributions of the parameters; in the event no correlation was detected and so the bivariate correlation was set to zero. The proposal variance for each parameter was set using trial and error, to keep the Metropolis acceptance ratio within the range 0.2-0.4. The bivariate proposal was implemented using the "mvrnorm" function ("MASS" R package, version 7.3-47; Venables and Ripley, 2002). The acceptance probability, *a*, in the Metropolis sampler was given by

$$\mathrm{a} = \min\left(1, exp\left[\ln\left(\pi(r^*_{max}, N^*_{2005})\right) - \ln\left(\pi(r_{max}, N_{2005})\right)\right]\right)$$

where $\pi(r_{max}, N_{2005})$ and $\pi(r^*_{max}, N^*_{2005})$ are the posterior densities evaluated at the current and proposed values of the parameters, respectively. In practice, both densities were computed up to a constant of proportionality, as described in the previous section.

A large number of samples to estimate posterior distributions was used, enabled by the relatively fast computation. Aiming to retain 10,000 samples per chain to approximate the posterior distributions of model parameters: 520,000 samples were simulated; thinning was applied by keeping every 50th sample, and discarding the rest to minimize storage; the 20,000 initial samples were discarded as burn-in. Three parallel chains were used to enable assessment of model convergence. Posterior distributions were calculated from the resulting 30,000 samples, 10,000 per chain.

Starting values for N_{2005} and r_{max} for each chain were chosen to be far enough apart to minimize the possibility of the Metropolis algorithm sampler misidentifying the most likely posterior distribution, i.e., preferentially sampling from a local maximum. Initial values for the three chains were 4,250, 6,250 and 8,250 for N_{2005}, and 0.01, 0.07 and 0.100 for r_{max}.

5.3 RESULTS

5.3.1 Base case model

Trace plots excluding burn-in samples indicated model convergence (Figure 5.4). The base case model implemented here produced very similar posteriors and population trajectory to those presented in Zerbini's base case (Table 5.3; compare Figure 5.3 with Figure 5.5 and Figure 5.1 with Figure 5.6). The base case model indicates that the current depletion of the WSA humpback whale population is 60% (N_{2019}: median = 14,552, 95% Credible Interval, CI = 7,282 - 19,874) of its carrying capacity (K: median = 24,524, 95% CI = 22,805 - 30,970).

Table 5.3. Posterior medians, means and, lower (2.5%) and upper (97.5%) limits of 95% credible (equal-tailed) interval of model parameters in Zerbini's base case scenario, the base case model and the updated model.

	Zerbini's base case			
Parameter	**Median**	**Mean**	**2.5%**	**97.5%**
r_{max}	0.069	0.066	0.013	0.104
K	24,558	25,110	22,791	31,118
N_{min}	503	850	159	3,943
Maximum depletion (year = 1958)	0.020	0.031	0.007	0.125
	Base case model			
	Median	**Mean**	**2.5%**	**97.5%**
r_{max}	0.069	0.066	0.014	0.104
K	24,524	25,060	22,805	30,970
N_{min}	496	843	233	3,809
Maximum depletion (year = 1958)	0.020	0.031	0.010	0.123
Current depletion (year = 2019)	0.597	0.581	0.240	0.860
	Updated model			
	Median	**Mean**	**2.5%**	**97.5%**
r_{max}	0.102	0.101	0.088	0.106
K	22,882	22,948	22,711	23,545
N_{min}	305	319	271	444
Maximum depletion (year = 1958)	0.013	0.014	0.012	0.019
Current depletion (year = 2019)	0.955	0.953	0.922	0.973

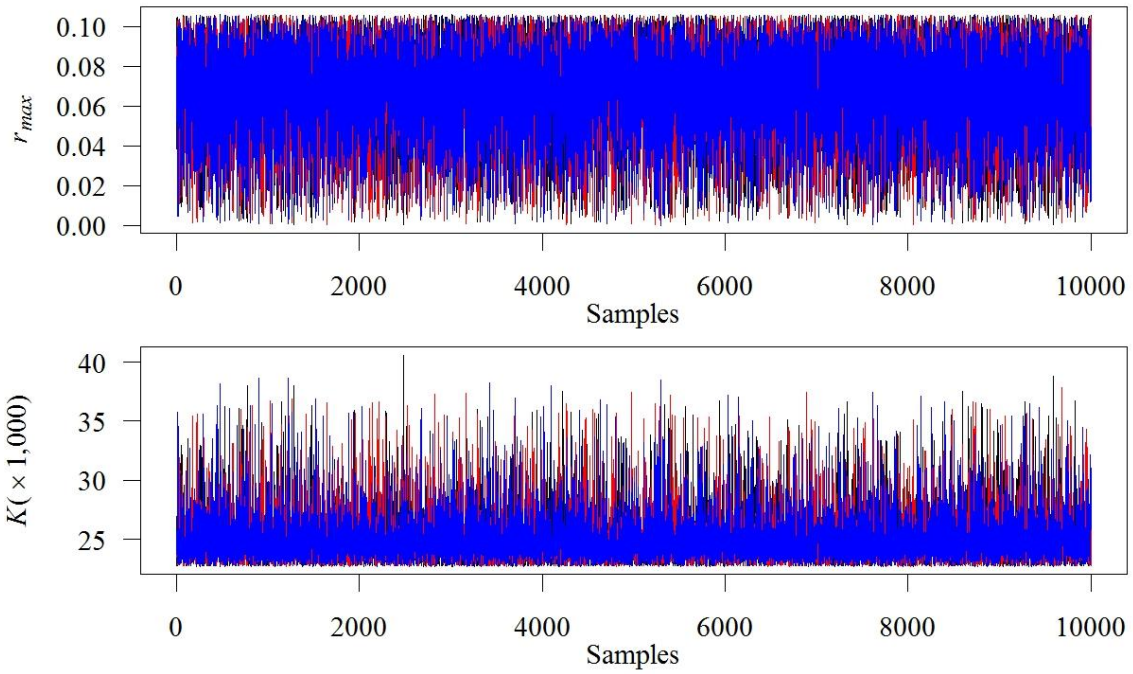


Figure 5.4. Base case model trace plots for K and r_{max} indicating model convergence, excluding burn-in. Three chains are represented by different colours (black, red and blue).

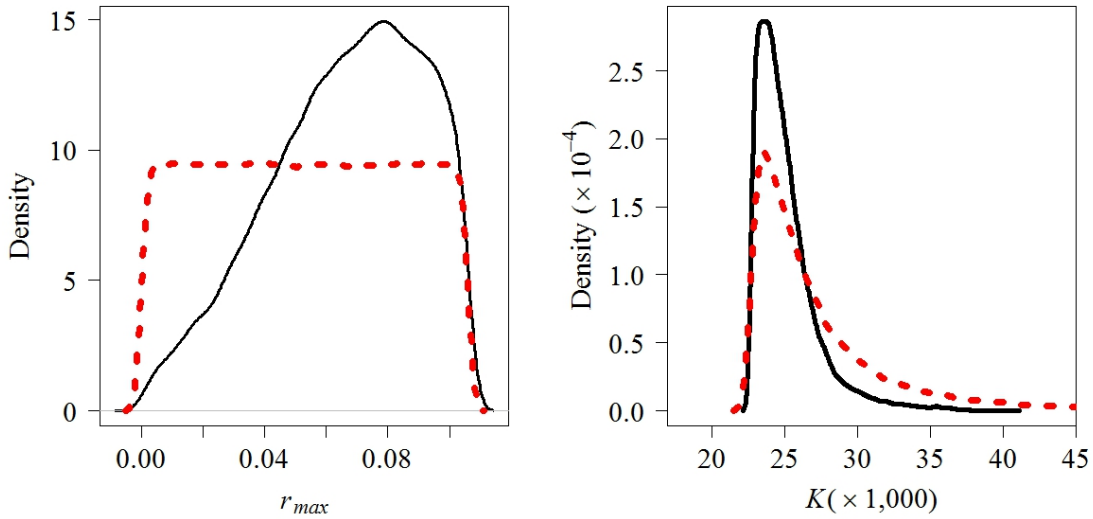


Figure 5.5. Posterior (continuous black lines) and prior (dashed red lines) distributions for r_{max} and K in the base case model. Compare this figure with Zerbini's base case scenario in Figure 5.3.

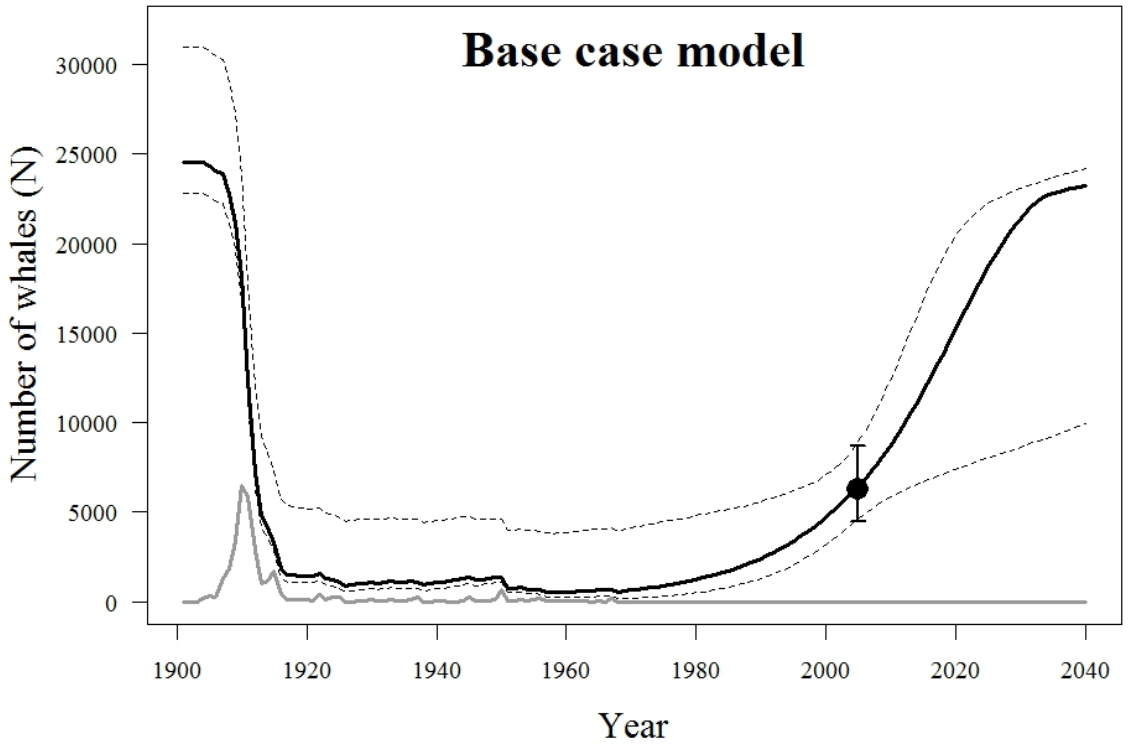


Figure 5.6. Population trajectory fitted for the base case model. Black lines represent the posterior median (solid line) and 95% credible interval limits (dashed lines) estimated for population size. Observed abundance in 2005 is indicated by a black dot with error bars (95% confidence interval). The grey line represents catches (Table 5.2). Compare this figure with Zerbini's base case scenario in Figure 5.1.

5.3.2 Updated model

Trace plots for this model after burn-in (Figure 5.7) again indicated model convergence. The updated model produced different posterior distributions from the base case model (Figure 5.8). For example, the carrying capacity was estimated to be about two thousand animals fewer (K: median = 22,882, 95% CI = 22,711-23,545) and the maximum recruitment rate was estimated to be about 50% higher (r_{max}: median = 0.102, 95% CI = 0.088-0.106). The estimated median maximum depletion indicated that the population size was as low as about 300 animals in 1958, almost 200 less than in the base case model. However, credible interval limits for the minimum estimated population size in the updated model (N_{min} 95% CI = 271 - 444) are completely within those from the base case model (N_{min} 95% CI = 233 - 3,809). Estimated current depletion indicates that the population is currently at about 95% of its carrying capacity, with a larger population size estimated for 2019 (N_{2019}: median = 21,878, 95% CI = 21,377-22,285) than in the base case model. Estimated abundances describing the population trajectory were also

more precise, with very narrow 95% credible intervals (Figure 5.9). The key piece of additional data causing this large increase in precision was further investigated in the "model checks" (see sections 5.22 and 5.33).

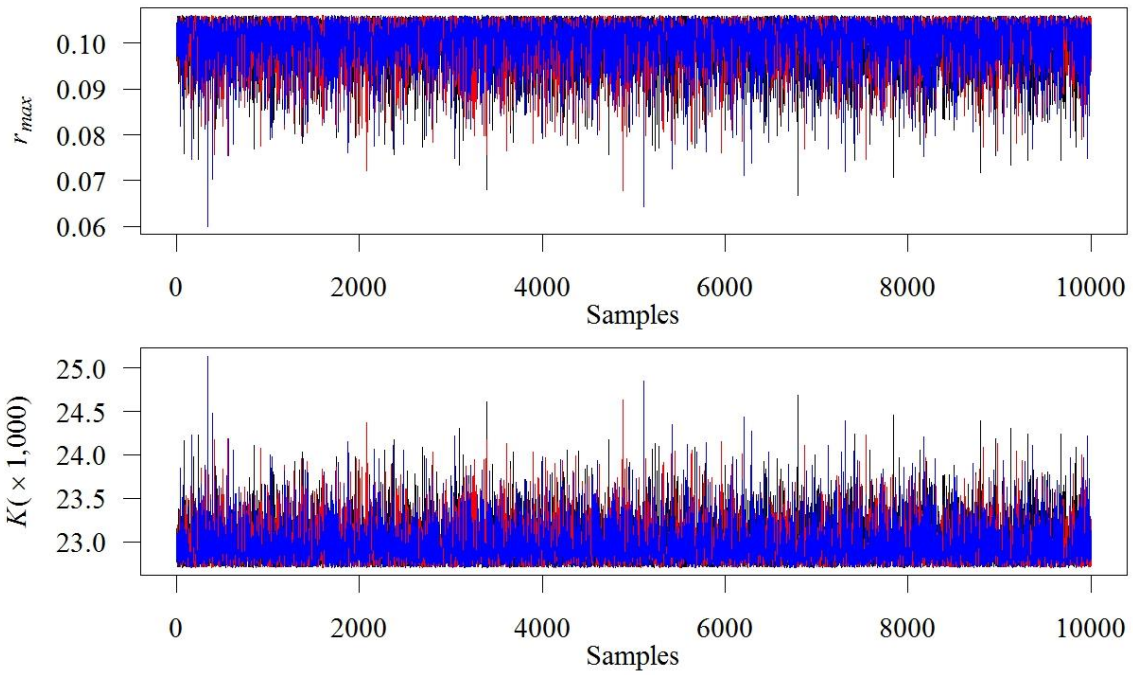


Figure 5.7. Updated model trace plots for K and r_{max} indicating model convergence, excluding burn-in. Three chains are represented by different colours (black, red and blue).

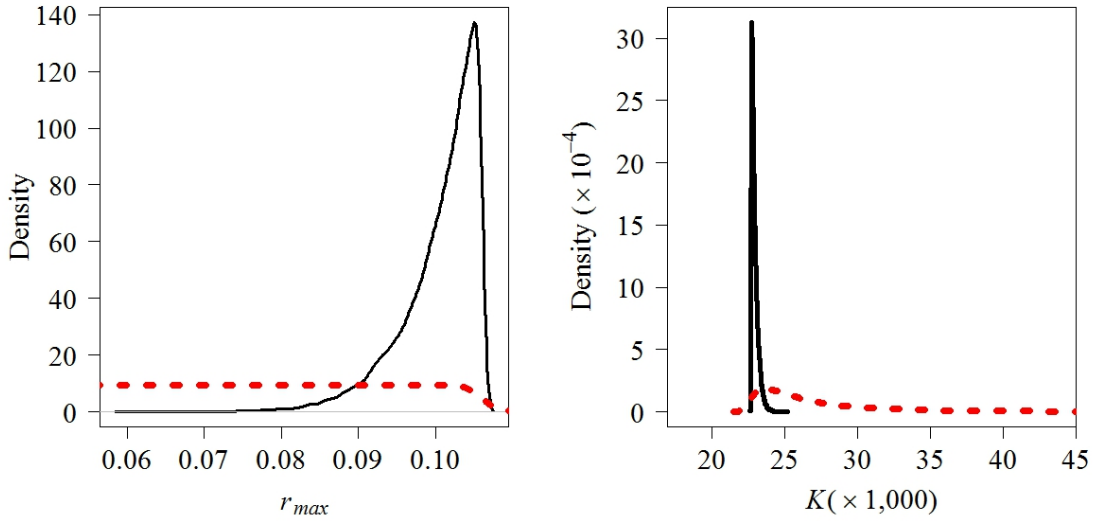


Figure 5.8. Posterior (continuous black lines) and prior (dashed red lines) distributions for r_{max} and K in the updated model. Horizontal axes are the same scale as Figure 5.5.

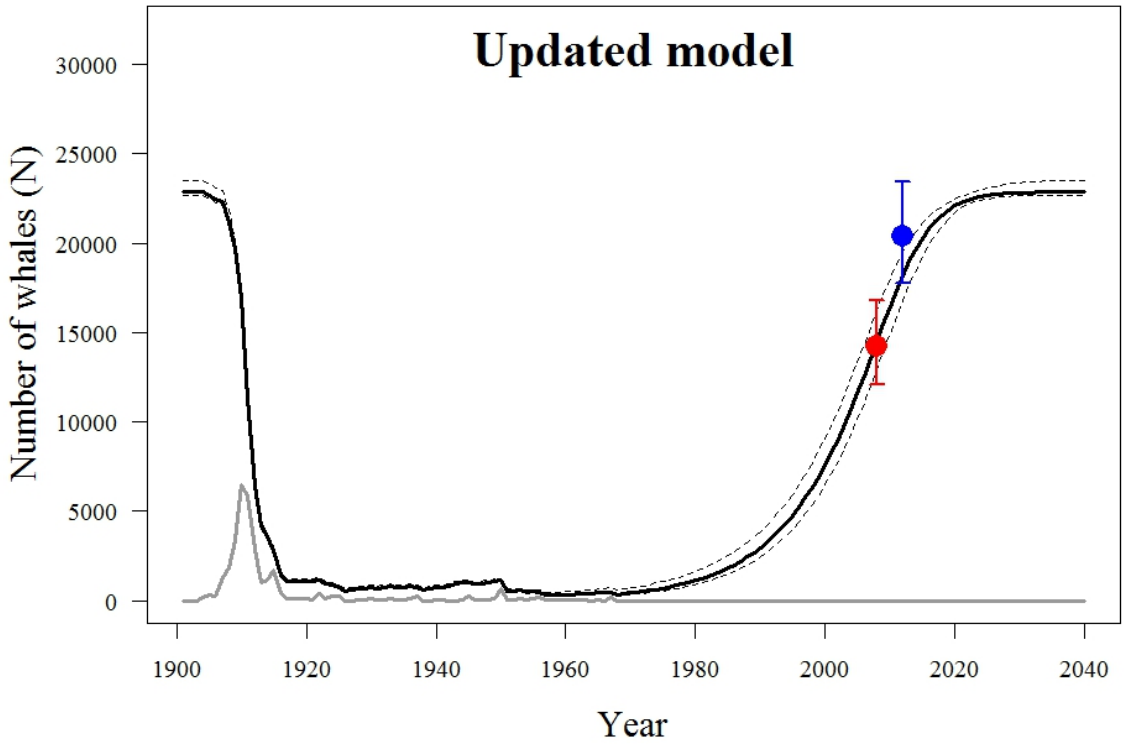


Figure 5.9. Population trajectory fitted for the updated model. Black lines represent the posterior median (solid line) and 95% credible interval limits (dashed lines) estimated for population size. Observed abundances in 2008 and 2012 are indicated, respectively, by red and blue dots with error bars (95% confidence intervals). The grey line represents catches (Table 5.2)

5.3.3 Model checks

Population trajectories estimated in the model check investigations (Figure 5.10) indicate that Model check 1, which was the only scenario to consider growth rate in 2002-2011, had relatively more precise estimates for population sizes. With the exception of this scenario and Model check 4, which included abundance estimates for both 2018 and 2012 (and excluded that for 2005), model checks had high uncertainty for *K*.

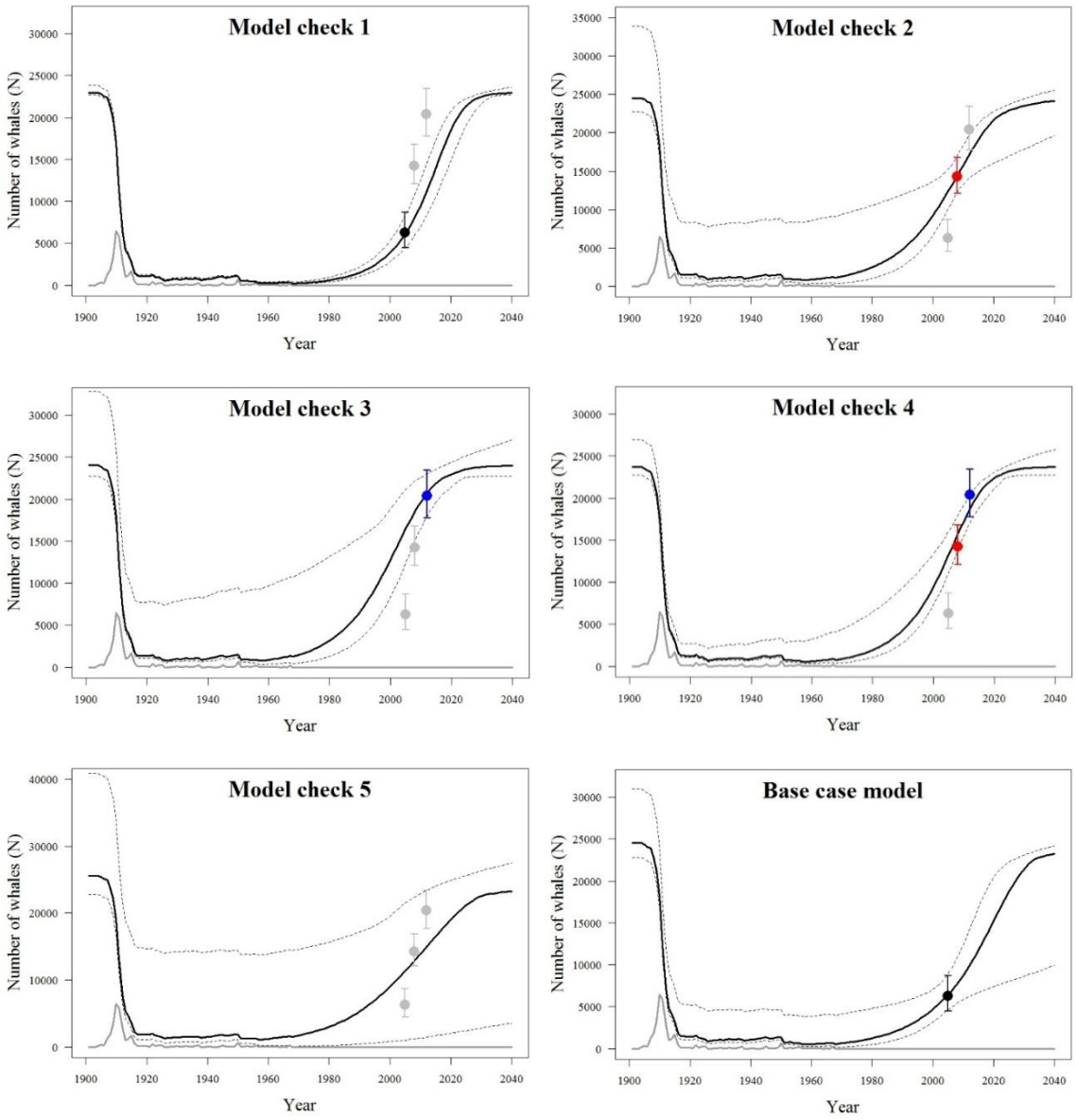


Figure 5.10. Population trajectory for the five Model checks and the base case model. Black lines represent the posterior median (solid line) and 95% credible interval limits (dashed lines) estimated for population size. Observed abundances in 2005, 2008 and 2012 are indicated by black, red and blue dots, respectively, with error bars (95% confidence intervals). Grey lines represent catches (Table 5.2). Grey dots with error bars indicate that the abundance data were not used in that Model check. Note the wider scale for the vertical axis in Model checks 2 and 5. Data inclusion in different Model checks are summarized in Table 5.4. The base case model population trajectory is presented for comparison.

Table 5.4. Summary of Model check outputs and data inclusion. Lower and upper 95% credible interval limits are indicated as 2.5% and 97.5%, respectively.

Model check	Parameter	Median	Mean	2.5%	97.5%	Data
1	r_{max}	0.100	0.098	0.082	0.106	Abundance for 2005, growth rate for 1995-1998 and growth rate for 2002-2011
	K	22969	23052	22715	23830	
	N_{min}	188	202	124	357	
	Max depletion	0.008	0.009	0.005	0.015	
2	r_{max}	0.070	0.067	0.012	0.104	Abundance for 2008 and growth rate for 1995-1998
	K	24498	25295	22795	33912	
	N_{min}	769	1565	289	8361	
	Max depletion	0.031	0.055	0.013	0.248	
3	r_{max}	0.078	0.074	0.018	0.105	Abundance for 2012 and growth rate for 1995-1998
	K	24033	24817	22764	32831	
	N_{min}	784	1392	529	7406	
	Max depletion	0.033	0.050	0.023	0.226	
4	r_{max}	0.085	0.081	0.040	0.105	Abundance for 2008 and 2012, and growth rate for 1995-1998
	K	23713	24006	22748	26942	
	N_{min}	516	789	291	3022	
	Max depletion	0.022	0.032	0.013	0.112	
5	r_{max}	0.054	0.053	0.003	0.103	No abundance or growth rate data
	K	25580	27136	22819	40935	
	N_{min}	1123	2691	257	13750	
	Max depletion	0.044	0.085	0.011	0.344	

5.4 DISCUSSION

The MCMC implemented here was able to produce very similar outputs to Zerbini's base case, when the same combination of data was considered, i.e., the base case model. This illustrates that the model implemented in that study can be replicated in a MCMC approach, providing independent verification of the model implementation. MCMC is a method with which ecologists are more familiar for estimating parameters in a Bayesian modelling framework (Kruschke, 2014), therefore the present implementation is more widely accessible. The inclusion of the new information on rate of increase for 2002-2011 and substitution of the observed abundance in 2005 by observed abundances in 2008 and 2012 resulted in changes in the estimated parameter posteriors and population trajectory.

The main difference in the updated model compared to the base case model is the higher precision in both the parameter posteriors and the population trajectory (Figure 5.5, Figure 5.6, Figure 5.8 and Figure 5.9; Table 5.3). Model checks allowed identification of the source of that improved precision. It is very clear from the population trajectory drawn for Model check 1 (Figure 5.10) that the growth rate information for 2002-2011 (Wedekin et al., 2017) is the main factor leading to improvement in precision. Growth rate for that nine-year period is relatively precise (CV = 0.115) and it represents a large portion of the data used for computing uncertainty (i.e., data observed with error). Model check 4, which excluded the abundance data for 2005 (Andriolo et al., 2006) but included estimated abundance for both 2008 and 2012 (Chapter 2), was the only other Model check to have an overall improvement in the population trajectory precision. The two abundance estimates are also very precise ($\hat{N}_{2008}$ CV = 0.084, $\hat{N}_{2012}$ CV = 0.071). Overall, the new growth rate and abundances estimates considered in the updated model seem to be consistent with each other. Another important point to highlight from the population trajectories (Figure 5.10) and posterior estimates (Table 5.4) from the Model checks, is that the posterior median for K was always estimated to be very close to its prior, although the credible interval is somewhat narrower when data are introduced. Model check 5, the scenario with no data considered other than the catch series, and the above observations show that the implied prior on K is quite informative (Figure 5.8), more so than one might wish given the high uncertainty about pre-exploitation population size, in the absence of more recent data.

While investigating Zerbini's model structure and specifications, some other potential updates and improvements for implementing the model became apparent. For example, the upper bound for the prior on r_{max} (0.106) was set based on a study that used a range of life history parameters from several humpback whale populations to compute the maximum plausible rate of population growth for the species (Clapham et al., 2006). However, a more recent publication by the same authors suggested the maximum plausible growth rate to be higher (0.118; [Zerbini et al., 2010]). Moreover, the estimates for r_{max} estimated from the updated model indicated most of its posterior density to be close to the upper bound imposed (i.e., 0.106). Additionally, the value of growth rate between 2002-2011 used in the present updated model (Wedekin et al., 2017) has a mean that is higher than that upper limit ($r_{2002\text{-}2011}$ mean = 0.1135, CV = 0.115), despite its 95%

credible interval ($r_{2002\text{-}2011}$ 95% CI = 0.088 - 0.139) including that value. An investigation into how to define a more appropriate prior on r_{max} should be considered in future modelling exercises.

Another potential topic of future investigation relates to the prior on carrying capacity adopted in Zerbini et al. (2011). The vague (i.e., not very informative) prior on N_{2005} combined with the vague prior on r_{max} leads to a rather informative prior on K (Figure 5.5) given the structure of the population model and catch data. For that reason, a vague prior on population size for any year in the base case would be expected to provide the same population trajectory and posteriors, so long as that prior was consistent with the implied value for K with the present prior on N_{2005}. In this sense, a potentially better approach to model the trajectory for this population, and to estimate the parameters of interest, would be to set a very vague prior directly on K. This, coupled with a more informed prior on r_{max} as described above, would likely provide important improvements on the present model implementation. This investigation is currently underway.

In the updated model, maximum depletion was estimated to be even more severe than before (Zerbini et al., 2011), with the posterior 95% probability credible interval indicating that the population could have been depleted to as low as 1% of its carrying capacity. That represents a value close to the minimum plausible size for this population, suggested to be 264 (Zerbini et al., 2011). This number was derived by applying a correction factor of four times (Jackson et al., 2006) the number of mtDNA haplotypes found in whales from this population (= 66; Rosenbaum et al., 2006), assuming an even sex ratio, and likely provides a conservative minimum bound for the population (Jackson et al., 2006). Although results from the updated model suggest that the median for the minimum population size was 305 (updated model N_{min}: median = 305, 95% CI = 271 - 444) in 1958, about 200 less than in the base case model (base case model N_{min}: median = 496, 95% CI = 233 - 3,809), this number still agrees very well with the above conservative genetic constraint information. Moreover, since carrying capacity was estimated to be lower in the updated model, with the posterior for maximum recruitment rate r_{max} indicating higher values to be more probable (updated model r_{max}: median = 0.102, 95% = 0.088-0.106) than in the base case model (base case model r_{max}: median = 0.069, 95% = 0.014-0.104), the population trajectory curve indicates that the

population could reach carrying capacity by around 2020 (Figure 5.9), sooner than indicated by Zerbini et al. (2011) and predicted with the base case model (i.e., at around 2040).

An important point to consider is that the models developed here assume that this population was at carrying capacity at the beginning of the period considered. Although there is information on pre-modern whaling catches, from before the 20th century (Morais et al., 2016), there is a lot of uncertainty in numbers caught for that period, which makes it difficult to include these data in the models. Many of the existing whaling records in the area are from small scale coastal whaling stations ("armações", in Portuguese), making it difficult to obtain whaling records. One possibility is to include a minimum number of whales caught for years pre-1901, which would help investigating a possible minimum number for carrying capacity. The inclusion of this information is likely to change drastically model outputs, with carrying capacity probably being estimated to be larger. Investigating different scenarios for catch series can help understanding the sensitivity of the model to the uncertainty on which scenario is the most realistic. Moreover, similarly to the observation processes described here for abundance and growth rate, a state-space model assuming that the catch data were observed with uncertainty can be fitted in a Bayesian framework. For that to be possible, measures of uncertainty for catch data are necessary, which currently do not exist.

Because only two reliable abundance estimates are available for this modelling exercise, ways to include more information on population size should be considered in future investigations. For example, although the 2005 abundance estimate from Andriolo et al (2006) may be biased low, it could potentially be included in the updated model if it were appropriately scaled up. That study used data collected in line transect aerial surveys to estimate WSA humpback whale population abundance and may have not adequately corrected for animals not detected on the trackline. Bortolotto et al. (2016a) discussed the differences in abundance estimates from line-transect ship survey data and from the line-transect aerial survey data used by Wedekin (2011) for the same year and using similar survey team members, aircraft and survey protocol as Andriolo et al. (2006). If the difference between abundances presented in Bortolotto et al. (2016a) and Wedekin (2011) can be used as a scaling factor to correct the abundance estimate from Andriolo et al.

(2006), the latter abundance estimate could possibly be included in the updated model presented here. Additionally, similar to the different scenarios of data inclusion investigated in Zerbini et al. (2011), other sources of estimated abundance could be considered to further update the present models. The recent study of Pavanato et al. (2017) is one such potential source: a Bayesian hierarchically structured model was applied to line transect data from aerial surveys; it also presented a compilation of abundance information published for this population in their breeding area off Brazil.

Models considered here assumed that carrying capacity is constant over time. Although this allows for a relatively simple modelling framework, this may not be realistic (Marshall and Quental, 2016). Also, none of the models allow for the possibility of non-natural mortalities, other than the whaling catches, which may currently occur (Bezamat et al., 2015; Ott et al., 2016). However, information on this is very sparse and there are no estimates of recent mortality due to human impacts for this population.

The results presented here further support that the WSA humpback whale population is increasing and will recover to its size prior to modern whaling in the next few years. However, future assessments of the conservation status of this population should consider how to incorporate information on pre-1901 whaling because this may lead to results that are different from those presented here.

APPENDIX 5.1

```
##PHD THESIS GUILHERME A. BORTOLOTTO
##CODE FOR IMPLEMENTING THE BASE CASE MODEL

##Bayesian population model, MCMC used for fitting model based on "Zerbini et
al. (2011) A Bayesian assessment of the conservation status of humpback whales
in the WSA"
##Last update 20/Jan/2019 by GAB

##This version sets a prior on number of animals in 2005 and uses backward
projection to derive an implied prior for number in the first year of inference
(1901)

library(parallel)
library(doParallel)
library(foreach)
library(doRNG)

#-------------------------------------------------------------------------------
# Function definitions
#-------------------------------------------------------------------------------

pop.trajectory <- function(rmax, K, catch, z = 2.39){
  ##Purpose: Create a population trajectory using the generalized logistic model
of Zerbini et al. (2011)
  # Inputs:
  # rmax - max population growth rate
  # K - carrying capacity (population assumed at carrying capacity at time 1)
  # catch - vector of catches
  # z - exponent in generalized logistic function
  ##Outputs:
  # vector of population sizes, of length equal to length of catch vector
  T <- length(catch)
  n <- numeric(T)
  n[1] <- K
  for(t in 1:(T-1)){
    # Equation 1 from Zerbini et al. (2011) paper
    n[t + 1] <- n[t] + n[t] * rmax * (1 - (max(n[t], 0) / K) ^ z) - catch[t]
  }
  return(n)
}

diff.k <- function(K, rmax, n.105, catch) {
  ##Purpose: Returns the difference between the projected population trajectory
in year 105 (i.e., 2005) given K, rmax and catch and the "known" population
value in that year, n.105.  This function is called repeatedly by find.k which
searches for the value of K that gives a diff closest to 0
  n <- pop.trajectory(rmax, K, catch)
  diff <- n[105] - n.105
  return(diff)
}

find.K <- function(rmax, n.105, catch){
  ##Purpose: Searches for a value of K (population size in year 1) given values
of rmax, n.105 ("known" number in 2005) and catch.  Uses bisection algorithm
implemented via uniroot function
  tol <- 1E-3
  res <- uniroot(diff.k, c(1, 100000), tol = tol, rmax, n.105, catch)
```

```
  return(res$root)
}

lnl <- function(n) {
  ##Purpose: return log-likelihood of pop trajectory n.
  # Assumes (1) normal observation of growth rate in 1995-98 and (2) lognormal
observation on population size in 2005
  # Inputs:
  # n - vector of population trajectory, starting in 1901
  ##Outputs:
  # log-likelihood as a scalar
  ##Log growth rate 1995.1998
  robs1995.1998 <- 0.074
  sigma1995.1998 <- 0.033
  ##Population size in 2005
  N2005 <- 6251
  CV.N2005 <- 0.17
  lN2005 <- log(N2005)
  sigma.lN2005 <- sqrt(log(1 + CV.N2005 ^ 2))
  ##Calculate observed log growth rate
  r1995.1998 <- (log(max(n[98], 0)) - log (max(n[95], 1E-7))) / 3
  ##lnl of observed rate
  res.1a <- -(log(sigma1995.1998) + 0.5 * (r1995.1998 - robs1995.1998) ^ 2 /
sigma1995.1998 ^ 2)
  ##lnl of population size in 2005
  res.2 <- -(log(sigma.lN2005) + lN2005 + 0.5 * (log(max(n[105], 0)) - lN2005) ^
2 / sigma.lN2005 ^ 2)

  res <- res.1a + res.2
  if (is.na(res)) res <- -Inf
  return(res)
}

ln.prior <- function(rmax, K, n.105) {
  ##Purpose: returns log of prior on rmax, K and n in 2005
  # Assume (1) uniform prior distribution for rmax (0, 0.106)
  # (2) uniform (500, 22000) prior on n in 2005
  ##Inputs:
  # rmax - value of rmax
  # n.105 - population size in 2005
  # K - value of K
  res.1 <- dunif(rmax, 0, 0.106, log = TRUE)
  res.2 <- dunif(n.105, 500, 22000, log = TRUE)
  res <- res.1 + res.2
  if(is.na(res)) res <- -Inf
  return(res)
}

generate.samples<-function(B, thin, starting.values, catch,
                           proposal.sd.mult = c(0.05, 0.05), proposal.rho = 0){
  ##Purpose: implements a random walk Metropolis sampler, based on a
multivariate normal proposal
  ##Inputs:
  # B - number of samples (before thinning)
  # thin - number of samples for each 1 saved
  # catch - vector of catches
  # starting.values - starting values (list) $rmax and $n.105
  # proposal.sd.mult - used for specifying the random walk sd - sd for each
parameter (vector length 2)
```

```
# is initial value * proposal.sd.mult
# proposal.rho - correlation between proposal for rmax and K
##Outputs:
# rmax - vector of B/thin rmax values
# n.105 - vector of B/thin N2005 values
# K - vector for B/thin K values
# n - matrix (B/thin * length(catch)) abundance values
# p.accept - proportion of proposed values that were accepted
##Implementation note - requires the MASS library
# for multivariate normal distribution
require(MASS)

##Set up storage of samples
n.to.save <- B%/%thin
rmax <- n.105 <- K <- numeric(n.to.save)
n <- matrix(0, n.to.save, length(catch))

##Initialize chain
rmax.current <- rmax[1] <- starting.values$rmax
n.105.current <- n.105[1] <- starting.values$n.105
K.current <- K[1] <- find.K(rmax.current, n.105.current, catch)
n.current <- n[1, ] <- pop.trajectory(rmax.current, K.current, catch)
ln.post <- lnl(n.current) + ln.prior(rmax.current, K.current, n.105.current)

##Random walk proposal specification
rmax.sigma.prop <- rmax.current * proposal.sd.mult [1]
n.105.sigma.prop <- n.105.current * proposal.sd.mult [2]
Sigma <- matrix(c(rmax.sigma.prop ^ 2,
                  rep(rmax.sigma.prop * n.105.sigma.prop * proposal.rho, 2),
                  n.105.sigma.prop ^ 2), 2, 2)

n.accept <- 0
for (i in 2:B){
  ##Propose new values
  prop <- mvrnorm(1, c(rmax.current, n.105.current), Sigma)
  rmax.prop <- prop[1]
  n.105.prop <- prop[2]
  ##Note: can't have zero or negative rmax or n.105
  if(rmax.prop > 0 & n.105.prop > 0) {
    K.prop <- find.K(rmax.prop, n.105.prop, catch)
    n.prop <- pop.trajectory(rmax.prop, K.prop, catch)
    ln.post.prop <- lnl(n.prop) + ln.prior(rmax.prop, K.prop, n.105.prop)

    acceptance.prob <- min(1, exp(ln.post.prop - ln.post))
    accept <- runif(1, 0, 1) < acceptance.prob
    if(accept) {
      n.accept <- n.accept + 1
      rmax.current <- rmax.prop
      n.105.current <- n.105.prop
      K.current <- K.prop
      n.current <- n.prop
      ln.post <- ln.post.prop
    }
  }

  ##Save samples
  if ((i-1) %% thin == 0) {
    s <- (i-1) %/% thin + 1
    rmax[s] <- rmax.current
```

```
      n.105[s] <- n.105.current
      K[s] <- K.current
      n[s, ] <- n.current
    }
  }

  p.accept <- n.accept / (B - 1)
  return(list(rmax = rmax, n.105 = n.105, K = K, n = n, p.accept = p.accept))
}

#-------------------------------------------------------------------------------
# Inputs
#-------------------------------------------------------------------------------

##Catch data
##Base catches from Zerbini et al., (2011) Table 4; Note: start in 1901
catch<-c(0, 0, 0, 180, 288, 240, 1261, 1849, 3391, 6468, 5832, 2881, 999,
         1155, 1697, 447, 121, 129, 111, 102, 9, 364, 133, 266, 254, 7, 0,
         19, 51, 107, 18, 23, 132, 57, 48, 105, 242, 0, 2, 36, 13, 0, 4, 60,
         238, 30, 35, 48, 83, 698, 45, 34, 140, 44, 96, 167, 61, 16, 15, 27,
         13, 24, 12, 0, 52, 0, 189, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0,
         0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
##project forward to 2040
catch <- c(catch, numeric(35))
T <- length(catch)
years <- 1:T + 1900

##MCMC chain specifications
n.chains <- 3
B <- 5.2E5
thin <- 50
n.to.save <- B%/%thin
n.saved<-n.to.save * n.chains
burnin <- 4e2
n.retain.samples <- n.to.save-burnin

##Starting values
starting.values<-data.frame(
  rmax = c(0.01, 0.07, 0.10),
  n.105 = c(4250, 6250, 8250))

##Comment this out to get a different result each time
set.seed(1234)

start.time <- Sys.time()  #Record time to run
if(n.chains>1) {
  ##Run multiple chains in parallel
  cl <- makePSOCKcluster(n.chains)
  registerDoParallel(cl)
  res <- foreach(chain = 1:n.chains) %dorng% {
    generate.samples(B = B, thin = thin, starting.values =
starting.values[chain,], catch, proposal.sd.mult = c(0.5, 0.5))
  }
  stopCluster(cl)
} else {
  ##Run a single chain without using parallelization
  res <- list(generate.samples(B = B, thin = thin, starting.values =
starting.values[1,], catch, proposal.sd.mult = c(0.5, 0.5)))
}
```

```
##Concatenate results from each chain
##Exclude burn-in samples: run only once, otherwise it will remove the burn-in
amount of samples repeatedly
for(i in 1:n.chains){
  res[[i]]$rmax <- res[[i]]$rmax[(burnin+1):length(res[[i]]$rmax)]
  res[[i]]$n.105 <- res[[i]]$n.105[(burnin+1):length(res[[i]]$n.105)]
  res[[i]]$K <- res[[i]]$K[(burnin+1):length(res[[i]]$K)]
  res[[i]]$n <- res[[i]]$n[(burnin+1):nrow(res[[i]]$n),]
}

str <- "c("
for(i in 1:n.chains){
  str <- paste(str, "res[[", i, "]]$rmax", ifelse(i < n.chains, ", ", ")"),
sep="")
}
rmax<-eval(parse(text=str))
str <- "c("
for(i in 1:n.chains){
  str <- paste(str, "res[[", i, "]]$n.105", ifelse(i < n.chains, ", ", ")"),
sep="")
}
n.105<-eval(parse(text=str))
str <- "c("
for(i in 1:n.chains){
  str <- paste(str, "res[[", i, "]]$K", ifelse(i < n.chains, ", ", ")"), sep="")
}
K<-eval(parse(text=str))
str <- "rbind("
for(i in 1:n.chains){
  str <- paste(str, "res[[", i, "]]$n", ifelse(i < n.chains, ", ", ")"), sep="")
}
n<-eval(parse(text=str))

n.median <- apply(n, 2, median)
n.lower <- apply(n, 2, quantile, 0.025)
n.upper <- apply(n, 2, quantile, 0.975)

end.time <- Sys.time()
time.diff <- end.time - start.time

## Save outputs
# save.image(file="Base case outputs") # commented out

#-------------------------------------------------------------------------------
# Results
#-------------------------------------------------------------------------------

##Plots
windows(record=TRUE, 10, 10)
par(mfrow = c(3, 1))

##Trace plots
plot(1:n.retain.samples, rep(0,n.retain.samples), type="n",
     ylim=range(rmax), xlab="sample", ylab="rmax")
for(i in 1:n.chains){
  lines(1:n.retain.samples, res[[i]]$rmax, col=i)
}
plot(1:n.retain.samples, rep(0,n.retain.samples), type="n",
```

```
     ylim=range(n.105), xlab="sample", ylab="n.105")
for(i in 1:n.chains){
  lines(1:n.retain.samples, res[[i]]$n.105, col=i)
}
plot(1:n.retain.samples, rep(0,n.retain.samples), type="n", ylim=range(K),
xlab="sample", ylab="K")
for(i in 1:n.chains){
  lines(1:n.retain.samples, res[[i]]$K, col=i)
}

##Posterior marginals
par(mfrow = c(2, 3))
hist(rmax); hist(n.105); hist(K)
plot(density(rmax)); plot(density(n.105)); plot(density(K))

##Joint posteriors
par(mfrow = c(1, 1))
plot(rmax, n.105, main=paste0("rmax x n.105 ", "(cor = ", round(cor(rmax,
n.105),4), ")"))
plot(rmax, K, main=paste0("rmax x K ", "(cor = ", round(cor(rmax, K),4), ")"))

##Plot of n
main = "Base case model"

##Add data points (N's)
##Log-normal CIs (as in p.77 in Buckland et al. (2001) - Introduction to
Distance Sampling)
N2005 <- 6251
CV.N2005 <- 0.17
point.N <- N2005
CV.N <- CV.N2005
varian.N <- (point.N*CV.N)^2
var.log.N <- log(1+(varian.N/(point.N^2)))
C <- exp(1.96*sqrt(var.log.N))
lower.N2005 <- N2005/C
upper.N2005 <- N2005*C

N2008 <- 14264
CV.N2008 <- 0.084
point.N <- N2008
CV.N <- CV.N2008
varian.N <- (point.N*CV.N)^2 #
var.log.N <- log(1+(varian.N/(point.N^2)))
C <- exp(1.96*sqrt(var.log.N)) #
lower.N2008 <- N2008/C
upper.N2008 <- N2008*C

N2012 <- 20389
CV.N2012 <- 0.071
point.N <- N2012
CV.N <- CV.N2012
varian.N <- (point.N*CV.N)^2
var.log.N <- log(1+(varian.N/(point.N^2)))
C <- exp(1.96*sqrt(var.log.N)) #
lower.N2012 <- N2012/C
upper.N2012 <- N2012*C

##Base case model population trajectory
windows(record=TRUE, 8,6)
```

```
par(mfrow=c(1,1), mar = c(5,5,4,2) + 0.1, family="serif") # default 'mar =
c(5,4,4,2) + 0.1'
main = "Base case model"
plot(years, n.median, ylim = c(0, 31000),
     type = "l", lwd = 3, main=main, las=1, ann=F)
title(ylab = "Number of whales (N)", line=3.6, cex.lab=1.5)
title(xlab = "Year", cex.lab=1.5)
title(main = main, cex.main=2, line=-2)
lines(years, n.lower, lty = 2, lwd = 1.5)
lines(years, n.upper, lty = 2, lwd = 1.5)
lines(years, catch, col = "grey60", lwd = 3)
points(2005, N2005, pch = 16, col = "black", cex = 2)
arrows(2005, lower.N2005, 2005, upper.N2005, length=0.05, angle=90, code=3,
lwd=2)

#-------------------------------------------------------------------------------
# END OF CODE
#-------------------------------------------------------------------------------
```

Intentionally left blank.

APPENDIX 5.2

```
##PHD THESIS GUILHERME A. BORTOLOTTO
##CODE FOR IMPLEMENTING THE UPDATED MODEL

##Bayesian population model, MCMC used for fitting model based on "Zerbini et
al. (2011) A Bayesian assessment of the conservation status of humpback whales
in the WSA"
##Last update 20/Jan/2019 by GAB

##This version sets a prior on number of animals in 2005 and uses backward
projection to derive an implied prior for number in the first year of inference
(1901)

library(parallel)
library(doParallel)
library(foreach)
library(doRNG)

#-------------------------------------------------------------------------------
# Function definitions
#-------------------------------------------------------------------------------

pop.trajectory <- function(rmax, K, catch, z = 2.39){
  ##Purpose: Create a population trajectory using the generalized logistic model
of Zerbini et al. (2011)
  ##Inputs:
  # rmax - max population growth rate
  # K - carrying capacity (population assumed at carrying capacity at time 1)
  # catch - vector of catches
  # z - exponent in generalized logistic function
  ##Outputs:
  # vector of population sizes, of length equal to length of catch vector
  T <- length(catch)
  n <- numeric(T)
  n[1] <- K
  for(t in 1:(T-1)){
    ##Equation 1 from Zerbini paper
    n[t + 1] <- n[t] + n[t] * rmax * (1 - (max(n[t], 0) / K) ^ z) - catch[t]
  }
  return(n)
}

diff.k <- function(K, rmax, n.105, catch) {
  #Purpose: Returns the difference between the projected population trajectory
in year 105 (i.e., 2005) given K, rmax and catch and the "known" population
value in that year, n.105.  This function is called repeatedly by find.k which
searches for the value of K that gives a diff closest to 0
  n <- pop.trajectory(rmax, K, catch)
  diff <- n[105] - n.105
  return(diff)
}

find.K <- function(rmax, n.105, catch){
  ##Purpose: Searches for a value of K (population size in year 1) given
  # values of rmax, n.105 ("known" number in 2005) and catch.  Uses bisection
algorithm implemented via uniroot function
  tol <- 1E-3
  res <- uniroot(diff.k, c(1, 100000), tol = tol, rmax, n.105, catch)
```

```
  #if(res$f.root > tol) stop ("Couldn't find K")
  return(res$root)
}

lnl <- function(n) {
  ##Purpose: return log-likelihood of pop trajectory n.
  # Assumes (1) normal observation of growth rate in 1995-98 and (2) lognormal
observation on population size in 2005
  ##Inputs:
  # n - vector of population trajectory, starting in 1901
  ##Outputs:
  # log-likelihood as a scalar
  ##Log growth rate 1995.1998
  robs1995.1998 <- 0.074
  sigma1995.1998 <- 0.033
  ##Log growth rate 2002.2011
  robs2002.2011 <- 0.1135
  sigma2002.2011 <- 0.013
  ##Population size in 2008 and 2012 (Chapter 2)
  N2008 <- 14264
  CV.N2008 <- 0.084
  lN2008 <- log(N2008)
  sigma.lN2008 <- sqrt(log(1 + CV.N2008 ^ 2))
  N2012 <- 20389
  CV.N2012 <- 0.071
  lN2012 <- log(N2012)
  sigma.lN2012 <- sqrt(log(1 + CV.N2012 ^ 2))
  ##Calculate observed log growth rate
  r1995.1998 <- (log(max(n[98], 0)) - log (max(n[95], 1E-7))) / 3
  r2002.2011 <- (log(max(n[111], 0)) - log (max(n[102], 1E-7))) / 9
  #lnl of observed rate
  res.1a <- -(log(sigma1995.1998) + 0.5 * (r1995.1998 - robs1995.1998) ^ 2 /
sigma1995.1998 ^ 2)
  res.1b <- -(log(sigma2002.2011) + 0.5 * (r2002.2011 - robs2002.2011) ^ 2 /
sigma2002.2011 ^ 2)
  #lnl of population size in 2005
  res.3 <- -(log(sigma.lN2008) + lN2008 + 0.5 * (log(max(n[108], 0)) - lN2008) ^
2 / sigma.lN2008 ^ 2)
  res.4 <- -(log(sigma.lN2012) + lN2012 + 0.5 * (log(max(n[112], 0)) - lN2012) ^
2 / sigma.lN2012 ^ 2)
  res <- res.1a + res.1b + res.3 + res.4
  if (is.na(res)) res <- -Inf
  return(res)
}

ln.prior <- function(rmax, K, n.105) {
  ##Purpose: returns log of prior on rmax, K and n in 2005
  # Assume (1) uniform prior distribution (0, 0.106) for rmax and
  #  (2) uniform (500, 22000) prior on n in 2005
  ##Inputs:
  # rmax - value of rmax
  # n.105 - population size in 2005
  # K - value of K
  res.1 <- dunif(rmax, 0, 0.106, log = TRUE)
  res.2 <- dunif(n.105, 500, 22000, log = TRUE)
  res <- res.1 + res.2
  if(is.na(res)) res <- -Inf
  return(res)
}
```

```
generate.samples<-function(B, thin, starting.values, catch,
                           proposal.sd.mult = c(0.05, 0.05), proposal.rho = 0){
  ##Purpose: implements a random walk Metropolis sampler, based on a
multivariate normal proposal
  ##Inputs:
  # B - number of samples (before thinning)
  # thin - number of samples for each 1 saved
  # catch - vector of catches
  # starting.values - starting values (list) $rmax and $n.105
  # proposal.sd.mult - used for specifying the random walk sd - sd for each
parameter (vector length 2)
  #   is initial value * proposal.sd.mult
  # proposal.rho - correlation between proposal for rmax and K
  ##Outputs:
  # rmax - vector of B/thin rmax values
  # n.105 - vector of B/thin N2005 values
  # K - vector for B/thin K values
  # n - matrix (B/thin * length(catch)) abundance values
  # p.accept - proportion of proposed values that were accepted
  ##Implementation note - requires the MASS library
  # for multivariate normal distribution
  require(MASS)
  ##Set up storage of samples
  n.to.save <- B%/%thin
  rmax <- n.105 <- K <- numeric(n.to.save)
  n <- matrix(0, n.to.save, length(catch))
  ##Initialize chain
  rmax.current <- rmax[1] <- starting.values$rmax
  n.105.current <- n.105[1] <- starting.values$n.105
  K.current <- K[1] <- find.K(rmax.current, n.105.current, catch)
  n.current <- n[1, ] <- pop.trajectory(rmax.current, K.current, catch)
  ln.post <- lnl(n.current) + ln.prior(rmax.current, K.current, n.105.current)

  ##Random walk proposal specification
  rmax.sigma.prop <- rmax.current * proposal.sd.mult [1]
  n.105.sigma.prop <- n.105.current * proposal.sd.mult [2]
  Sigma <- matrix(c(rmax.sigma.prop ^ 2,
                    rep(rmax.sigma.prop * n.105.sigma.prop * proposal.rho, 2),
                    n.105.sigma.prop ^ 2), 2, 2)
  n.accept <- 0
  for (i in 2:B){
    ##Propose new values
    prop <- mvrnorm(1, c(rmax.current, n.105.current), Sigma)
    rmax.prop <- prop[1]
    n.105.prop <- prop[2]
    ##Can't have zero or negative rmax or n.105
    if(rmax.prop > 0 & n.105.prop > 0) {
      K.prop <- find.K(rmax.prop, n.105.prop, catch)
      n.prop <- pop.trajectory(rmax.prop, K.prop, catch)
      ln.post.prop <- lnl(n.prop) + ln.prior(rmax.prop, K.prop, n.105.prop)

      acceptance.prob <- min(1, exp(ln.post.prop - ln.post))
      accept <- runif(1, 0, 1) < acceptance.prob
      if(accept) {
        n.accept <- n.accept + 1
        rmax.current <- rmax.prop
        n.105.current <- n.105.prop
        K.current <- K.prop
```

```
        n.current <- n.prop
        ln.post <- ln.post.prop
      }
    }
    ##Save samples
    if ((i-1) %% thin == 0) {
      s <- (i-1) %/% thin + 1
      rmax[s] <- rmax.current
      n.105[s] <- n.105.current
      K[s] <- K.current
      n[s, ] <- n.current
    }
  }
  p.accept <- n.accept / (B - 1)
  return(list(rmax = rmax, n.105 = n.105, K = K, n = n, p.accept = p.accept))
}

#-------------------------------------------------------------------------------
# Inputs
#-------------------------------------------------------------------------------

##Catch data
##Base catches from Zerbini et al., (2011) Table 4; Note: start in 1901
catch<-c(0, 0, 0, 180, 288, 240, 1261, 1849, 3391, 6468, 5832, 2881, 999,
         1155, 1697, 447, 121, 129, 111, 102, 9, 364, 133, 266, 254, 7, 0,
         19, 51, 107, 18, 23, 132, 57, 48, 105, 242, 0, 2, 36, 13, 0, 4, 60,
         238, 30, 35, 48, 83, 698, 45, 34, 140, 44, 96, 167, 61, 16, 15, 27,
         13, 24, 12, 0, 52, 0, 189, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0,
         0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)

##Project forward to 2040
catch <- c(catch, numeric(35))
T <- length(catch)
years <- 1:T + 1900

##MCMC chain specifications
n.chains <- 3
B <- 5.2E5
thin <- 50
n.to.save <- B%/%thin
n.saved<-n.to.save * n.chains
burnin <- 4e2
n.retain.samples <- n.to.save-burnin

##Starting values
starting.values<-data.frame(
  rmax = c(0.01, 0.07, 0.10),
  n.105 = c(4250, 6250, 8250))

#Comment this out to get a different result each time
set.seed(1234)

start.time <- Sys.time()  #Record time to run
if(n.chains>1) {
  ##Run multiple chains in parallel using foreach construct
  cl <- makePSOCKcluster(n.chains)
  registerDoParallel(cl)
  res <- foreach(chain = 1:n.chains)  %dorng% {
```

```
    generate.samples(B = B, thin = thin, starting.values =
starting.values[chain,], catch, proposal.sd.mult = c(0.5, 0.5))
  }
  stopCluster(cl)
} else {
  ##Run a single chain without using parallelization
  res <- list(generate.samples(B = B, thin = thin, starting.values =
starting.values[1,], catch, proposal.sd.mult = c(0.5, 0.5)))
}

##Concatenate results from each chain
for(i in 1:n.chains){
  res[[i]]$rmax <- res[[i]]$rmax[(burnin+1):length(res[[i]]$rmax)]
  res[[i]]$n.105 <- res[[i]]$n.105[(burnin+1):length(res[[i]]$n.105)]
  res[[i]]$K <- res[[i]]$K[(burnin+1):length(res[[i]]$K)]
  res[[i]]$n <- res[[i]]$n[(burnin+1):nrow(res[[i]]$n),]
}

str <- "c("
for(i in 1:n.chains){
  str <- paste(str, "res[[", i, "]]$rmax", ifelse(i < n.chains, ", ", ")"),
sep="")
}
rmax<-eval(parse(text=str))
str <- "c("
for(i in 1:n.chains){
  str <- paste(str, "res[[", i, "]]$n.105", ifelse(i < n.chains, ", ", ")"),
sep="")
}
n.105<-eval(parse(text=str))
str <- "c("
for(i in 1:n.chains){
  str <- paste(str, "res[[", i, "]]$K", ifelse(i < n.chains, ", ", ")"), sep="")
}
K<-eval(parse(text=str))
str <- "rbind("
for(i in 1:n.chains){
  str <- paste(str, "res[[", i, "]]$n", ifelse(i < n.chains, ", ", ")"), sep="")
}
n<-eval(parse(text=str))

n.median <- apply(n, 2, median)
n.lower <- apply(n, 2, quantile, 0.025)
n.upper <- apply(n, 2, quantile, 0.975)

end.time <- Sys.time()
time.diff <- end.time - start.time

##Save outputs
#save.image(file="Updated_model_outputs")

#-------------------------------------------------------------------------------
# Results
#-------------------------------------------------------------------------------

##Plots
windows(record=TRUE, 10, 10)

##Trace plots
```

```
par(mfrow = c(3, 1))
plot(1:n.retain.samples, rep(0,n.retain.samples), type="n",
     ylim=range(rmax), xlab="sample", ylab="rmax")
for(i in 1:n.chains){
  lines(1:n.retain.samples, res[[i]]$rmax, col=i)
}
plot(1:n.retain.samples, rep(0,n.retain.samples), type="n",
     ylim=range(n.105), xlab="sample", ylab="n.105")
for(i in 1:n.chains){
  lines(1:n.retain.samples, res[[i]]$n.105, col=i)
}
plot(1:n.retain.samples, rep(0,n.retain.samples), type="n", ylim=range(K),
xlab="sample", ylab="K")
for(i in 1:n.chains){
  lines(1:n.retain.samples, res[[i]]$K, col=i)
}
##Posterior marginals
par(mfrow = c(2, 3))
hist(rmax); hist(n.105); hist(K)
plot(density(rmax)); plot(density(n.105)); plot(density(K))

##Joint posteriors
par(mfrow = c(1, 1))
plot(rmax, n.105, main=paste0("rmax x n.105 ", "(cor = ", round(cor(rmax,
n.105),4), ")"))
plot(rmax, K, main=paste0("rmax x K ", "(cor = ", round(cor(rmax, K),4), ")"))

##Plot of n
main = "Updated model"

##Add data points (N's)
##Log-normal CIs (as in p.77 in Buckland et al. (2001) - Introduction to
Distance Sampling)
N2005 <- 6251
CV.N2005 <- 0.17
point.N <- N2005
CV.N <- CV.N2005
varian.N <- (point.N*CV.N)^2
var.log.N <- log(1+(varian.N/(point.N^2)))
C <- exp(1.96*sqrt(var.log.N))
lower.N2005 <- N2005/C
upper.N2005 <- N2005*C

N2008 <- 14264
CV.N2008 <- 0.084
point.N <- N2008
CV.N <- CV.N2008
varian.N <- (point.N*CV.N)^2 #
var.log.N <- log(1+(varian.N/(point.N^2)))
C <- exp(1.96*sqrt(var.log.N)) #
lower.N2008 <- N2008/C
upper.N2008 <- N2008*C

N2012 <- 20389
CV.N2012 <- 0.071
point.N <- N2012
CV.N <- CV.N2012
varian.N <- (point.N*CV.N)^2
var.log.N <- log(1+(varian.N/(point.N^2)))
```

```
C <- exp(1.96*sqrt(var.log.N)) #
lower.N2012 <- N2012/C
upper.N2012 <- N2012*C

#Updated model population trajectory
windows(record=TRUE, 8,6)
par(mfrow=c(1,1), mar = c(5,5,4,2) + 0.1, family="serif") # default 'mar =
c(5,4,4,2) + 0.1'
main = "Updated model"
plot(years, n.median, ylim = c(0, 30000),
     type = "l", lwd = 3, main=main, las=1, ann=F)
title(ylab = "Number of whales (N)", line=3.6, cex.lab=1.5)
title(xlab = "Year", cex.lab=1.5)
title(main = main, cex.main=2, line=-2)
lines(years, n.lower, lty = 2, lwd = 1.5)
lines(years, n.upper, lty = 2, lwd = 1.5)
lines(years, catch, col = "grey60", lwd = 3)

points(2008, N2008, pch = 16, cex = 2, col="red")
arrows(2008, lower.N2008, 2008, upper.N2008, length=0.05, angle=90, code=3,
lwd=2, col="red")
points(2012, N2012, pch = 16, cex = 2, col="blue")
arrows(2012, lower.N2012, 2012, upper.N2012, length=0.05, angle=90, code=3,
lwd=2, col="blue")

#------------------------------------------------------------------------------
# END OF CODE
#------------------------------------------------------------------------------
```

Intentionally left blank.

CHAPTER 6

GENERAL DISCUSSION

6.1 THESIS SUMMARY

This thesis was motivated by the lack of knowledge on how the population of humpback whales uses space and habitats in Brazilian coastal waters and, more generally, how migratory whales relate to and are influenced by the environment in their breeding areas. It was also motivated by the need for information about human impacts on that population, which is one of the most successful in terms of post-whaling recovery, growing at the upper limit of the rate plausibly expected for the species (Wedekin et al., 2017). Information on the distribution of animal populations and how they use their habitats can be used to evaluate the need for, and to support implementation of, management actions intended to improve conservation status.

Humpback whales typically migrate every year between summer feeding grounds in high latitudes, and winter breeding grounds in tropical and sub-tropical regions (Clapham, 2000). Migration is "an adaptation to resources that fluctuate spatiotemporally either seasonally or less predictably" and it "plays a central role in the spatial dynamics of mobile populations" (Dingle and Drake, 2007). Therefore, for animals that migrate between distinct breeding and feeding areas, environmental characteristics or resources related their use of space are expected to differ: in feeding areas, the main resource is food/prey availability, while in breeding areas, the resources of interest are those related to special conditions for breeding (Dingle, 2014).

Reasons for conserving migratory whales include that they provide important ecological services to their environments. For example, when humpback whales migrate they move nutrients from highly productive waters in high latitude feeding grounds, where they prey on small crustaceans and fish, to the usually less-productive waters in their breeding grounds (Roman et al., 2014). Whales provide important nutrients to primary producers in these habitats in the form of their faeces or from their carcasses, when they die.

Therefore, they take part in multiple trophic roles in both ends of their migration (Dolphin, 1987; Thomas et al., 1991; Morissette et al., 2006).

In this thesis, distribution and habitat use of humpback whales off the coast of Brazil were modelled using two different data types, which provided new information on area use by these animals during the breeding season. New estimates of abundance derived from these models provided a new perspective on the increase in the population. To facilitate future assessment of the impacts of the developing oil and gas industry, a simple method for investigating the potential risk to the population from oil spills was developed, making use of the new results on distribution. To put these studies in a historical perspective and to update information on conservation status, a population dynamics model that has previously been used to assess status and inform management decisions was revised and updated with the new estimates of abundance, and with other available data.

6.2 DISTRIBUTION AND HABITAT USE

Distribution and habitat use of a population result from animals seeking resources in space to meet their life requirements, from the effects of the environment and from their physiological restrictions (Aarts and Matthiopoulos, 2010). Thus, in the breeding grounds of migratory whales, resources that improve chances of animals reproducing are likely the primary drivers of distribution (Dingle, 2014). While in feeding areas the influence of prey distribution on how animals use space is more clear, habitat use in breeding grounds is mainly influenced by social, behavioural and other individual characteristics, which are not as clearly understood.

In that sense, different individual-level characteristics, such as age, sex, reproductive stage, social role, among others things, can influence the distribution and habitat use of cetaceans in breeding areas (Craig and Herman, 2000; Ersts and Rosenbaum, 2003; Elwen and Best, 2004a; b; Oviedo and Solís, 2008; Félix and Botero-Acosta, 2011; Cartwright et al., 2012; Craig et al., 2014; Rayment et al., 2015). The distribution models fitted in Chapters 2 and 3 did not include individual characteristics, and conclusions were only possible at the population level. Future work to investigate this may provide important additional information on how humpback whales use their breeding habitat.

The whale telemetry data used in Chapter 3 could be used to investigate the potential influence of environmental features on how animals use space from a behavioural perspective: statistical models can be applied to investigate relationships between environmental features and behaviour states, which are inferred from movement patterns (Jonsen, 2016). Investigating different behaviour states within breeding areas, along migration routes and in feeding areas could provide insights into different habitat requirements for animals in these distinct regions. Because information on sex, social role and group composition are available in the telemetry data used here (Table 3.1), hierarchical models in the framework of Jonsen (2016) could be fitted to these data to infer about habitat use for specific cohorts.

Preliminary work to apply such models to the Brazilian humpback whale telemetry data has been initiated, which has identified differences in movement patterns between when animals were migrating and when they were in other areas. Area-restricted search was observed along the coast of Brazil and also in waters around South Georgia and the South Sandwich Islands. The latter result supports the previous identification (Stevick et al., 2006; Zerbini et al., 2006; Engel and Martin, 2009) of these sub-Antarctic islands as feeding grounds.

The main result from investigating habitat use for this population in their breeding grounds was the strong relationship between humpback whale density/occurrence and sea surface temperature (SST), estimated from line transect data modelling (Chapter 2) and telemetry tracking data modelling (Chapter 3). It is not surprising that SST plays an important role in how whales use their habitat, because they have evolved anatomical/physiological adaptions to enable them to live their entire lives in an environment where heat loss is much greater than experienced by their terrestrial predecessors (Gingerich, 2015). This strong relationship has a major influence on the distribution patterns predicted by the models (Figures 2.11, 3.4 and 3.5) and, therefore, on the assessment on potential risk from oil spills (Chapter 4). However, it has been suggested that, because this population is increasing, animals are redistributing and reoccupying areas used before intensive whaling, when abundance was higher (Rossi-Santos et al., 2008; Gonçalves et al, 2018). If animals are indeed changing their

distribution as a result of the increase in the number of animals, this may be better verified in the near future when the population is expected to reach carrying capacity (Chapter 5).

A possible reoccupation of areas may be a result of the increase in population size (Chapter 5). For example, humpback whales were known to breed off waters of the Dominican Republic and along the Antillean chain before commercial whaling in the area (Whitehead and Moore, 1982). Due to the extensive exploitation of animals, whales became mostly absent along the Antillean arc Islands, including Guadeloupe, despite of the region being historically considered an important whaling ground (Smith et al., 1999). After cessation of whaling, populations in the north Atlantic are recovering and Stevick et al. (2016) recently reported individuals recorded in both breeding grounds off Cape Verde Islands, Africa, and off Guadeloupe. No animal was re-sighted between Cape Verde and the Dominican Republic (nor between Guadalupe and the Dominican Republic), despite the photograph catalog used from the latter region being about ten times larges than that from Guadeloupe. That study suggested that there may be distinct population groups in the West Indies, with whales found in Cape Verde interbreeding with those found in Guadalupe. The presence of whales in the latter coincides with the increase in North Atlantic humpback whale population sizes after protection against whaling (Stevick et al., 2003). However, the large numbers of humpback whales once found in Cape Verde are not seen today, indicating that the area has not been repopulated after the end of whaling (Wenzel et al., 2009). Another example is the great increase in records of north-eastern Pacific grey whales calving in the north of Mexico, outside the traditional "calving lagoons" of Baja California, Mexico (Shelden et al., 2004). This also coincides with the recovery in population size after whaling (Buckland and Breiwick, 2002).

The importance of SST found here for these animals adds support to the hypothesis that large whale migration is motivated by improving the chances of calf survival, because calves may benefit from specific temperature requirements to conserve energy (Corkeron and Connor, 2001). An improvement to the spatial models would be to investigate additional or different covariates that could be related to shelter, not only to explore habitat use further but also to understand better the meaning of shelter to whales. It is logical that whales in breeding areas are likely to prefer areas where ideal shelter

conditions are available (Chapter 2), however what measurable environmental features are part of that, remains unclear. For example, Rayment et al. (2015) considered sheltered waters in right whale breeding grounds off Port Ross, in the north of the sub-Antarctic Auckland Islands, to be areas that were protected from exposure to prevailing wind (i.e., westerly winds). That study indicated that calving female southern right whales seek shallower waters, closer to the shore, and sheltered from wind and also sheltered form wave exposure. Contrasting to the present habitat use models (Chapters 2 and 3) in which shelter included the effect of SST, water temperature was not investigated in Rayment et al. (2015) as its effects on right whale distribution were considered negligible *a priori* by those authors, due to the low temperatures and very narrow range found in that habitat (6.1-7.7 °C). It is very possible that what constitutes shelter is different for different migratory whales species, therefore different species may favour different habitats in the breeding grounds.

Distribution patterns estimated here confirm previous observations and studies showing that humpback whales concentrate around the Abrolhos Archipelago (Siciliano, 1997; Martins et al., 2001; Andriolo et al., 2010). The new distribution information presented (Chapters 2 and 3) is a larger scale quantitative description of areas used by humpback whales in Brazilian waters. However, animals have been found beyond the limits considered here (Lodi, 1994; Wedekin et al., 2014; Bortolotto et al., 2016c) and obtaining data in a wider area could improve future investigations. For example, humpback whales have been found around the Trindade Islands (Siciliano et al., 2012; Wedekin et al., 2014; Lucena et al., 2015), about 1,000 km away from the closest point on the continental shelf and Townsend (1935) had reported two animals caught in that area in the nineteenth century. Whether those areas around oceanic islands represent important breeding grounds for the population, or are simply a passing point for migrating animals, could be investigated with analysis of additional dedicated survey effort.

6.3 CONSERVATION AND MANAGEMENT

Although most populations of large whales around the world are recovering after intensive commercial whaling (Clapham et al., 1999; Stevick et al., 2003; Zerbini et al., 2011; Laake et al., 2012; George et al., 2015), animals are now facing pressures from a

range of other human activities (Halpern et al., 2008). Among them is the exploration of oil in the marine environment.

Because oil exploration activities at sea began, and were intensified over the years, in the more accessible areas closer to the coast, migratory whales when in coastal breeding areas may be at higher risk of impact from such activities than when in offshore waters. Animals exposed to oil may get intoxicated through ingestion, inhalation or contact with the eyes and mouth mucosae (Geraci and St Aubin, 1990), with calves likely being at higher risk because of their smaller body size and less developed immune system, being more susceptible to being impacted. In whale breeding areas such as the Brazilian coast, the increasing interest in oil exploration may represent an important conservation concern in the future.

The work described in Chapter 4 on the risk of oil spill impacts on the Brazilian humpback whale population, which used results from Chapters 2 and 3, was partially motivated to improve similar investigations presented in previously published studies on risk of impact. For example, Martins et al. (2013) overlapped humpback whale densities estimated within survey blocks in Brazil (Andriolo et al., 2010) with marine areas where oil exploration activities were happening or planned to suggest areas that should be prioritized for conservation of humpback whales in Brazil. Although the discussion of what sort of conservation actions may be needed is another point to be considered, as discussed below, that study did not account for two important aspects investigated in Chapter 4: the potential impact of oil spills, which are not spatially fixed along the breeding season, therefore the dynamic nature of spills must be considered; and the finer resolution of whale distribution. Additionally, the method presented here can be easily adapted to identifying areas of higher risk of impact from oil spills to other marine species, given that a distribution surface is available. Therefore, this can support management decisions, such as restricting oil exploration in certain areas, to reduce the risks of impact to populations of communities in the marine environment. In summary, management of human activities can be better informed with updated information on potential impacts, allowing a better understanding of the population's conservation status.

The importance of investigating the status of populations is on permitting animals to reach and maintain good conservation status. To evaluate that, population assessments require

some specific information depending on the expected impacts. For example, in the case of unintended impact, if conservation status is pessimist, then threats need to be evaluated to determine if management actions are required and, what type of actions should be taken. When the impact is caused through whaling, this process is necessary to assess the direct impact of exploitation on population status to manage, for example, sustainable whaling quotas. Therefore, the International Whaling Commission (IWC) has promoted population assessments for whale species to investigate the two above points (e.g., IWC, 2002; 2003). Methods currently used to assess whale populations status evolved from models used in fisheries biology, and adapted by the IWC from 1960 (Wade, 2018). Challenges faced by the IWC in such assessments included understanding population structure, or stock identity, because different populations overlap their distribution, especially in their feeding grounds (IWC, 2002), causing uncertainties associated with which catches in feeding grounds should be associated with each population. As an example, the "Sub-Committee on the Comprehensive Assessment of North Atlantic Humpback Whales" of the IWC Scientific Committee met in 2001 (IWC, 2002), and again in 2002 (IWC, 2003), to review information on stock identity, catches, abundance, trends and other biological parameters to improve assessments for North Atlantic humpback whale populations. The method used for assessment at the time was a density-dependent, age- and sex-structured population dynamics model fitted to data. The usefulness of such assessments dependents on the availability of the best and most up-to-date information possibly available.

The population assessment investigated here was based on a simpler model, i.e., age- and sex- aggregated (Zerbini et al., 2011), because more detailed information was not available, given that studies of the Brazilian humpback whale population are relatively recent compared to North Atlantic populations. Although many improvements can potentially be made, the main issue with the model investigated here was that the prior implied for carrying capacity, K, was too informative. As Chapter 5 shows, the abundance and growth rate data used in those models, which were observed with error, do not contribute much for estimating K, given the population model structure and catch data, assumed to have been observed without error. The priors on maximum recruitment, r_{max}, and in population size for 2005 practically define K, meaning that not much is "learned"

by the model from the data, with priors having the ultimate influence on posteriors (Figures 5.3 and 5.5).

To investigate this further, work has been initiated on a modelling exercise that considers an explicit vague prior on *K* (i.e., not implied through backward projection), a better informed prior on r_{max}, and no priors on population size other than on *K* (population size at the beginning of the period considered), to investigate a potentially more realistic estimation of the population carrying capacity. There have been investigations and discussions in the literature (Zerbini et al., 2010; Clapham and Zerbini, 2015, Wedekin et al., 2017) presenting information on the growth rate of the population, which can contribute to a better informed prior on maximum recruitment rate, r_{max}. Given that this was one of the only two parameters to have priors explicitly specified in the state-space models from Chapter 5, this could greatly further benefit investigations. Another important point to consider in future investigations of population status is to obtain and be able to incorporate updated information about catches (Morais et al., 2017). The fact that animals were hunted along the coast of Brazil before 1900, even if not at such a large scale as subsequently, may mean that the currently estimated population size in 1901 is not realistic, and may not represent the carrying capacity of the population, *K*.

Abundance estimates presented in Chapter 2 are very precise (CV = 0.084 for 2008; CV = 0.071 for 2012), which is unusual for cetacean studies, including those applying density surface models to estimate abundance (e.g., Roberts et al., 2016; Kanaji et al., 2017; Cañadas et al., 2018). However, encounter rates in studies for rare species are inevitably lower and more variable, causing final abundances to be less precise (Cañadas et al., 2018), which did not happen in the present surveys, since whales were abundant. In design-based abundance estimates, the majority of final variance is typically from encounter rate and group size estimations (Buckland et al., 2015). In model-based approaches, such as those in Chapter 2, the fitted relationships to covariates can explain part of that variance, resulting in more precise estimates. The reasons for the precision in estimates presented here are likely a combination of factors, which include the very precise detection function (CV = 0.044) and relatively high level of explanatory power of the abundance estimation model (abundance estimation model, AEM, explained deviance = 66.8%). Also, group sizes were corrected or confirmed when resightings were

possible and when in high-density areas, with off-effort observers helping to track groups, avoid double-counting and with species identification, but not with searching nor detecting new animals. That procedure likely improved precision in the data, and possibly also improved accuracy. In summary, the more accurate and precise line transect data available, the more accurate and precise will be evaluations of population conservation status (Chapter 5).

6.4 CONCLUDING REMARKS

Findings from this thesis support the following general conclusions:

- The distribution of migratory animals in breeding areas can be strongly related to environmental conditions that enhance calf survival and development, as indicated by the spatial modelling results for SST and shelter covariates;
- Modelling of distribution and habitat use from line transect and individual tracking data can be complementary, but several differences in the nature of the data and methods used must be carefully considered. Other modelling approaches, especially those that allow individual information to be specified, have the potential to further complement such investigations;
- The proposed oil dispersion simulation is a fast and easy way for identification of areas where oil spills present higher risk of impacting whales, conditional on animal distribution maps and data on ocean currents being available. However, more investigations are needed to validate the method as useful;
- Assessments of population status through Bayesian state-space population dynamics models can be greatly improved with additional data. Careful consideration of model structure and prior assumptions are extremely important;
- The humpback whale population breeding off Brazil will likely soon reach its size prior to whaling in the twentieth century and its conservation status is therefore good. However, the animals are under increasing pressure in their breeding grounds as human activities expand in these areas.

Intentionally left blank.

REFERENCES

Aarts, G., MacKenzie, M., McConnell, B., Fedak, M. & Matthiopoulos, J. (2008) Estimating space-use and habitat preference from wildlife telemetry data. *Ecography*. 31 (1), 140–160.

Adcroft, A., Hallberg, R., Dunne, J.P., Samuels, B.L., Galt, J.A., Barker, C.H. & Payton, D. (2010) Simulations of underwater plumes of dissolved oil in the Gulf of Mexico. *Geophysical Research Letters*. 37 (18), 1–5.

Allaire, J.J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J. & Chang, W. (2017) *rmarkdown: dynamic documents for R. R package version 1.8*. Available from: https://cran.r-project.org/package=rmarkdown.

Allison, C. (2006) Documentation of the creation of the Southern Hemisphere humpback catch series. SC/A06/HW47. *IWC Workshop on Comprehensive Assessment of Southern Hemisphere Humpback Whales*. 9.

Alonso, M.B., Feo, M.L., Corcellas, C., Vidal, L.G., Bertozzi, C.P., Marigo, J., Secchi, E.R., Bassoi, M., Azevedo, A.F., Dorneles, P.R., Torres, J.P.M., Lailson-Brito, J., Malm, O., Eljarrat, E. & Barceló, D. (2012) Pyrethroids: a new threat to marine mammals? *Environment International*. 47, 99–106.

Alves, L.C.P.D.S., Moreira, S., Simões-lopes, P.C. & Andriolo, A. (2010) Behavioral responses of humpback whales, *Megaptera novaeangliae* (Cetacea: Balaenopteridae), to satellite transmitter deployment procedures. *Zoologia*. 27 (1), 1–6.

Amante, C. & Eakins, B.W. (2009) *ETOPO1 1 Arc-Minute Global Relief Model: procedures, data sources and analysis. NOAA Technical Memorandum NESDIS NGDC-24. National Geophysical Data Center, NOAA. doi:10.7289/V5C8276M*.

Amr, T. & Stamboliyska, R. (2016) 'Getting the facts right' in Tarek Amr & Rayna Stamboliyska (eds.) *Practical D3.js*. Paris: Apress. pp. 39–59.

Andrews, R.D., Pitman, R.L. & Ballance, L.T. (2008) Satellite tracking reveals distinct movement patterns for Type B and Type C killer whales in the southern Ross Sea, Antarctica. *Polar Biology*. 31, 1461–1468.

Andriolo, A., Kinas, P.G., Engel, M.H., Martins, C.C.A. & Rufino, A.M. (2010) Humpback whales within the Brazilian breeding ground: distribution and population size estimate. *Endangered Species Research*. 11 (3), 233–243.

Andriolo, A., Martins, C.C.A., Engel, M.H., Pizzorno, J.L., Más-Rosa, S., Freitas, A.C., Morete, M.E. & Kinas, P.G. (2006) The first aerial survey to estimate abundance of humpback whales (*Megaptera novaeangliae*) in the breeding ground off Brazil (Breeding Stock A). *Journal of Cetacean Research and Management*. 8 (3), 307–311.

ANP (2018a) *ANP conclui as investigações do acidente no Campo de Frade. http://www.anp.gov.br/noticias/1545-anp-conclui-as-investigacoes-do-acidente-no-campo-de-frade*. Available from: http://www.anp.gov.br/noticias/1545-anp-conclui-as-investigacoes-do-acidente-no-campo-de-frade (*accessed 4 December 2018*).

ANP (2018b) *BDEP webmaps.* Available from: http://www.anp.gov.br/exploracao-e-producao-de-oleo-e-gas/dados-tecnicos?view=default (*accessed 5 September 2018*).

Bailey, H., Corkrey, R., Cheney, B. & Thompson, P.M. (2013) Analyzing temporally correlated dolphin sightings data using generalized estimating equations. *Marine Mammal Science*. 29 (1), 123–141.

Bailey, H., Hammond, P.S. & Thompson, P.M. (2014) Modelling harbour seal habitat by combining data from multiple tracking systems. *Journal of Experimental Marine Biology and Ecology*. 450, 30–39.

Bailey, H. & Thompson, P.M. (2009) Using marine mammal habitat modelling to identify priority conservation zones within a marine protected area. *Marine Ecology Progress Series*. 378, 279–287.

Baracho-Neto, C.G., Santos Neto, E., Rossi-Santos, M.R., Wedekin, L.L., Neves, M.C., Lima, F. & Faria, D. (2012) Site fidelity and residence times of humpback whales

(*Megaptera novaeangliae*) on the Brazilian coast. *Journal of the Marine Biological Association of the United Kingdom*. 92 (8), 1–9.

Baumgartner, M.F. & Mate, B.R. (2005) Summer and fall habitat of North Atlantic right whales (*Eubalaena glacialis*) inferred from satellite telemetry. *Canadian Journal of Fisheries and Aquatic Sciences*. 62 (3), 527–543.

Bezamat, C., Wedekin, L.L. & Simões-Lopes, P.C. (2015) Potential ship strikes and density of humpback whales in the Abrolhos Bank breeding ground, Brazil. *Aquatic Conservation: Marine and Freshwater Ecosystems*. 25, 712–725.

Bivand, R.S., Pebesma, E. & Gomez-Rubio, V. (2013) *Applied spatial data analysis with R*. 2nd edition. Springer, NY.

Bivand, R. & Rundel, C. (2017) *rgeos: interface to Geometry Engine - Open Source ('GEOS'). R package version 0.3-26.*

Block, B.A., Jonsen, I.D., Jorgensen, S.J., Winship, A.J., Shaffer, S.A., Bograd, S.J., Hazen, E.L., Foley, D.G., Breed, G.A., Harrison, A., Ganong, J.E., Swithenbank, A., Castleton, M., Dewar, H., Mate, B.R., Shillinger, G.L., Schaefer, K.M., Benson, S.R., Weise, M.J., et al. (2011) Tracking apex marine predator movements in a dynamic ocean. *Nature*. 475 (7354), 86–90.

Bortolotto, G.A., Danilewicz, D., Andriolo, A., Secchi, E.R. & Zerbini, A.N. (2016a) Whale, whale, everywhere: increasing abundance of western South Atlantic humpback whales (*Megaptera novaeangliae*) in their wintering grounds. *PLoS ONE*. 11 (10), e0164596.

Bortolotto, G.A., Danilewicz, D., Andriolo, A. & Zerbini, A.N. (2016b) Humpback whale *Megaptera novaeangliae* (Cetartiodactyla: Balaenopteridae) group sizes in line transect ship surveys: An evaluation of observer errors. *Zoologia*. 33 (2), 1–5.

Bortolotto, G.A., Kolesnikovas, C.K.M., Freire, A.S. & Simões-Lopes, P.C. (2016c) Young humpback whale *Megaptera novaeangliae* feeding off Santa Catarina coastal waters, Southern Brazil, and a ship strike report. *Marine Biodiversity Records*. 9 (29).

Bortolotto, G.A., Morais, I.O.B., Reis, M.S.S. & Souto, L.R.A. (2016d) Anthropogenic impact on a pregnant Cuvier's beaked whale *Ziphius cavirostris* stranded in Brazil. *Marine Biodiversity Records*. 9 (30).

Bortolotto, G.A., Danilewicz, D., Hammond, P.S., Thomas, L. & Zerbini, A.N. (2017) Whale distribution in a breeding area: spatial models of habitat use and abundance of western South Atlantic humpback whales. *Marine Ecology Progress Series*. 585, 213–227.

Bradshaw, C.J.A., Higgins, J., Michael, K.J., Wotherspoon, S.J. & Hindell, M.A. (2004) At-sea distribution of female southern elephant seals relative to variation in ocean surface properties. *ICES Journal of Marine Science*. 61 (6), 1014–1027.

BRASIL (2015) *Portal Brasil*. Available from: http://www.brasil.gov.br/ (*accessed 9 July 2015*).

Brekke, C. & Solberg, A.H.S. (2005) Oil spill detection by satellite remote sensing. *Remote Sensing of Environment*. 95 (1), 1–13.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. & Thomas, L. (2001) *Introduction to distance sampling*. Oxford: Oxford University Press.

Buckland, S.T. & Breiwick, J.M. (2002) Estimated trends in abundance of eastern Pacific gray whales from shore counts (1967/68 to 1995/96). *Journal of Cetacean Research and Management*. 4 (1), 41–48.

Buckland, S.T., Rexstad, E.A., Marques, T.A. & Oedekoven, C.S. (2015) *Distance sampling: Methods and applications*. *Methods in statistical ecology*. Andrew P. Robinson, Stephen T. Buckland, Peter Reich, & Michael McCarthy (eds.). Springer International Publishing.

Butterworth, D.S. & Best, P.B. (1994) The origins of the choice of 54% of carrying capacity as the protection level for baleen whale stocks, and the implications thereof for management procedures. SC/F92/Mg2. *Report of the International Whaling Commission 44*.

Butterworth, D.S. & Punt, A.E. (1995) On the Bayesian approach suggested for the assessment of the Bering-Chukchi-Beaufort Seas stock of bowhead whales. SC/46/AS2. *Report of the International Whaling Commission*. 45, 303–311.

Cañadas, A., Aguilar de Soto, N., Aissi, M., Arcangeli, A., Azzolin, M., B-Nagy, A., Bearzi, G., Frantzis, A., Garcia, P., Gazo, M., Gutierrez-Xarxa, R., Holcer, D., Laran, S., Lauriano, G., Lewisv, T., Moulinsw, A., Mussi, B., Notarbartolo di Sciara, G., Panigada, S., et al. (2018) The challenge of habitat modelling for threatened low density species using heterogeneous data: the case of Cuvier's beaked whales in the Mediterranean. *Ecological Indicators*. 85, 128–136.

Cañadas, A., Sagarminaga, R., De Stephanis, R., Urquiola, E. & Hammond, P.S. (2005) Habitat preference modelling as a conservation tool: proposals for marine protected areas for cetaceans in southern Spanish waters. *Aquatic Conservation: Marine and Freshwater Ecosystems*. 15 (5), 495–521.

Cartwright, R., Gillespie, B., LaBonte, K., Mangold, T., Venema, A., Eden, K. & Sullivan, M. (2012) Between a rock and a hard place: habitat selection in female-calf humpback whale (*Megaptera novaeangliae*) pairs on the Hawaiian breeding grounds. *PLoS ONE*. 7 (5), e38004.

Cassoff, R.M., Moore, K.M., Mclellan, W.A., Barco, S.G., Rotstein, D.S. & Moore, M.J. (2011) Lethal entanglement in baleen whales. *Diseases of Aquatic Organisms*. 96175–185.

Castro, F.R., Mamede, N., Danilewicz, D., Geyer, Y., Pizzorno, J.L.A., Zerbini, A.N. & Andriolo, A. (2014) Are marine protected areas and priority areas for conservation representative of humpback whale breeding habitats in the western South Atlantic? *Biological Conservation*. 179, 106–114.

Cerchio, S., Trudelle, L., Zerbini, A.N., Charrassin, J., Geyer, Y., Mayer, F.X., Jung, J., Adam, O. & Rosenbaum, H.C. (2016) Satellite telemetry of humpback whales off Madagascar reveals insights on breeding behavior and long-range movements within the southwest Indian Ocean. *Marine Ecology Progress Series*. 562, 193–209.

Clapham, P., Wade, P. & Zerbini, A. (2006) Plausible rates of population growth in humpback whales revisited. SC/58/SH4. May 2006, St. Kitts and Nevis, West Indies. *Report of the IWC Scientific Committee*. 12pp.

Clapham, P.J. (2000) The humpback whale: seasonal feeding and breeding in a baleen whale, in J. Mann (ed.) *Cetacean societies: field studies of dolphins and whales*. Chicago: University of Chicago Press. pp. 173–196.

Clapham, P.J., Aguilar, A. & Hatch, L.T. (2008) Determining spatial and temporal scales for management: lessons from whaling. *Marine Mammal Science*. 24 (1), 183–201.

Clapham, P.J. & Baker, C.S. (2018) 'Whaling, modern', in *Encyclopedia of Marine Mammals*. 3rd edition. pp. 1070–1074.

Clapham, P.J., Young, S. & Brownell Jr., R. (1999) Baleen whales: conservation issues and the status of the most endangered populations. *Mammal Review*. 29 (1), 35–60.

Clapham, P.J. & Zerbini, A.N. (2015) Are social aggregation and temporary immigration driving high rates of increase in some Southern Hemisphere humpback whale populations? *Marine Biology*. 1–12.

Cooke, A. (2018) *Megaptera novaeangliae, humpback whale*. Available from: http://dx.doi.org/10.2305/IUCN.UK.2018-2.RLTS.T13006A50362794.en (*accessed 20 January 2019*).

Corkeron, P.J. & Connor, R.C. (1999) Why do baleen whales migrate? *Marine Mammal Science*. 15 (4), 1228–1245.

Craig, A.S. & Herman, L.M. (2000) Habitat preferences of female humpback whales *Megaptera novaeangliae* in the Hawaiian Islands are associated with reproductive status. *Marine Ecology Progress Series*. 193, 209–216.

Craig, A.S., Herman, L.M., Pack, A.A. & Waterman, J.O. (2014) Habitat segregation by female humpback whales in Hawaiian waters: avoidance of males? *Behaviour*. 151, 613–631.

Crain, C.M., Halpern, B.S., Beck, M.W. & Kappel, C.V. (2009) Understanding and managing human threats to the coastal marine environment. *Annals of the New York Academy of Sciences*. 1162, 39–62.

Croll, D.A., Clark, C.W., Calambokidis, J., Ellison, W.T. & Tershy, B.R. (2001) Effect of anthropogenic low-frequency noise on the foraging ecology of *Balaenoptera* whales. *Animal Conservation*. 4, 13–27.

Cui, D. & Zhao, P. (2018) Tracking buoys research for oil spilling with the wireless charging equipment. *AIP Conference Proceedings* 1944, 020070.

Dähne, M., Gilles, A., Lucke, K., Peschko, V., Adler, S., Krügel, K., Sundermeyer, J. & Siebert, U. (2013) Effects of pile-driving on harbour porpoises (*Phocoena phocoena*) at the first offshore wind farm in Germany. *Environmental Research Letters*. 8 (2), 025002.

Dalla Rosa, L., Ford, J.K.B. & Trites, A.W. (2012) Distribution and relative abundance of humpback whales in relation to environmental variables in coastal British Columbia and adjacent waters. *Continental Shelf Research*. 36, 89–104.

Danilewicz, D., Moreno, I.B., Ott, P.H., Tavares, M., Azevedo, A.F., Secchi, E.R. & Andriolo, A. (2010) Abundance estimate for a threatened population of franciscana dolphins in southern coastal Brazil: uncertainties and management implications. *Journal of the Marine Biological Association of the United Kingdom*. 90 (08), 1649–1657.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., Van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., et al. (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*. 137, 553–597.

Dias, L.A., Litz, J., Garrison, L., Martinez, A., Barry, K. & Speakman, T. (2017) Exposure of cetaceans to petroleum products following the Deepwater Horizon oil spill in the Gulf of Mexico. *Endangered Species Research*. 33 (1), 119–125.

Dingle, H. (2014) *Migration: the biology of life on the move*. 2nd edition. Oxford University Press.

Dingle, H. & Drake, V.A. (2007) What is migration? *BioScience*. 57 (2), 113–121.

Dolphin, W.F. (1987) Prey densities and foraging of humpback whales, *Megaptera novaeangliae*. *Experientia*. 43 (4), 468–471.

Dorneles, P.R., Lailson-Brito, J., Secchi, E.R., Dirtu, A.C., Weijs, L., Dalla Rosa, L., Bassoi, M., Cunha, H.A., Azevedo, A.F. & Covaci, A. (2015) Levels and profiles of chlorinated and brominated contaminants in Southern Hemisphere humpback whales, *Megaptera novaeangliae*. *Environmental Research*. 138, 49–57.

Dunlop, R.A., Noad, M.J., Mccauley, R.D., Scott-hayward, L., Kniest, E., Slade, R., Paton, D. & Cato, D.H. (2017) Determining the behavioural dose-response relationship of marine mammals to air gun noise and source proximity. *Journal of Experimental Biology*. 220, 2878–2886.

Edmundson, W. & Hart, I. (2014) *A história da caça de baleias no Brasil: de peixe real a iguaria Japonesa*. São Paulo: DISAL.

Elwen, S.H. & Best, P.B. (2004a) Environmental factors influencing the distribution of southern right whales (*Eubalaena australis*) on the south coast of South Africa I: broad scale patterns. *Marine Mammal Science*. 20 (3), 567–582.

Elwen, S.H. & Best, P.B. (2004b) Female southern right whales *Eubalaena australis*: are there reproductive benefits associated with their coastal distribution off South Africa? *Marine Ecology Progress Series*. 269, 289–295.

Embling, C.B., Gillibrand, P.A., Gordon, J., Shrimpton, J., Stevick, P.T. & Hammond, P.S. (2010) Using habitat models to identify suitable sites for marine protected areas for harbour porpoises (*Phocoena phocoena*). *Biological Conservation*. 143 (2), 267–279.

Engel, M.H., Marcondes, M.C.C., Martins, C.C.A. & Luna, F.O. (2004) *Are seismic surveys responsible for cetacean strandings? An unusual mortality of adult humpback*

whales in Abrolhos Bank, Northeastern coast of Brazil. Work SC/56/E28 presented to the International Whaling Commission Scientific Committee, Sorrento, Italy.

Engel, M.H. & Martin, A.R. (2009) Feeding grounds of the western South Atlantic humpback whale population. *Marine Mammal Science*. 25 (4), 964–969.

Erbe, C. (2002) Underwater noise of whale-watching boats and potential effects on killer whales (*Orcinus orca*), based on an acoustic impact model. *Marine Mammal Science*. 18 (2), 394–418.

Ersts, P.J. & Rosenbaum, H.C. (2003) Habitat preference reflects social organization of humpback whales (*Megaptera novaeangliae*) on a wintering ground. *Journal of Zoology*. 260337–345.

ESR (2009) *OSCAR third degree resolution ocean surface currents, PO.DAAC., Ver. 1*. Available from: podaac.jpl.nasa.gov/dataset/OSCAR_L4_OC_third-deg *(accessed in 04 Apr 2016).*

Félix, F. & Botero-Acosta, N. (2011) Distribution and behaviour of humpback whale mother-calf pairs during the breeding season off Ecuador. *Marine Ecology Progress Series*. 426, 277–287.

Fernandez, E. & Lellouche, J.M. (2018) *Product user manual for the Global Ocean Physical Reanalysis product GLOBAL_REANALYSIS_ PHY_001_030*. (1.1).

Fieberg, J., Rieger, R.H., Zicus, M.C. & Schildcrout, J.S. (2009) Regression modelling of correlated data in ecology: subject-specific and population averaged response patterns. *Journal of Applied Ecology*. 46 (5), 1018–1025.

Findlay, K.P., Seakamela, S.M., Mey, M.A., Kirkman, S.P., Barendse, J., Cade, D.E., Hurwitz, D., Kennedy, A.S., Kotze, P.G.H., Mccue, S.A., Thornton, M., Vargas-Fonseca, O.A. & Wilke, G. (2017) Humpback whale 'super-groups': a novel low-latitude feeding behaviour of Southern Hemisphere humpback whales (*Megaptera novaeangliae*) in the Benguela upwelling system. *PLoS ONE.* 12(3): e01720021.

Forcada, J. (2004) 'Distribution', in William F. Perrin, Bernd Würsig, & J.G.M. Thewissen (eds.) *Encyclopedia of Marine Mammals*. 2nd edition. Academic Press. pp. 316–321.

Francini-Filho, R.B. & Moura, R.L. (2008) Dynamics of fish assemblages on coral reefs subjected to different management regimes in the Abrolhos Bank, eastern Brazil. *Aquatic Conservation: Marine and Freshwater Ecosystems*. 18, 1166–1179.

Fraser, G.S. (2014) 'Impacts of offshore oil and gas development on marine wildlife resources', in J. Gates, D. Trauger, & B. Czech (eds.) *Peak oil, economic growth, and wildlife conservation*. New York, NY: Springer. pp. 191–217.

Freitas, A.C., Kinas, P.G., Martins, C.C.A. & Engel, M.H. (2004) Abundance of humpback whales on the Abrolhos Bank wintering ground, Brazil. *Journal of Cetacean Research and Management*. 6 (3), 225–230.

Friedlaender, A.S., Halpin, P.N., Qian, S.S., Lawson, G.L., Wiebe, P.H., Thiele, D. & Read, A.J. (2006) Whale distribution in relation to prey abundance and oceanographic processes in shelf waters of the Western Antarctic Peninsula. *Marine Ecology Progress Series*. 317, 297–310.

Garrigue, C., Zerbini, A.N., Geyer, Y., Heide-Jørgensen, M.-P., Hanaoka, W. & Clapham, P. (2010) Movements of satellite-monitored humpback whales from New Caledonia. *Journal of Mammalogy*, 91 (1):109–115.

George, J.C., Druckenmiller, M.L., Laidre, K.L., Suydam, R. & Person, B. (2015) Bowhead whale body condition and links to summer sea ice and upwelling in the Beaufort Sea. *Progress in Oceanography*. 136, 250–262.

Geraci, J.R. & St. Aubin, D.J. (1990) *Sea mammals and oil: confronting the risks*. San Diego, CA: Academic Press.

Geraci, J.R., St Aubin, D.J. & Reisman, R.J. (1983) Bottlenose dolphins, *Tursiops truncatus*, can detect oil. *Canadian Journal of Fisheries and Aquatic Sciences*. 40, 1516–1522.

Geyer, C.J. (1992) Practical Markov Chain Monte Carlo. *Statistical Science*. 7 (4), 473–511.

Gilles, A., Scheidat, M. & Siebert, U. (2009) Seasonal distribution of harbour porpoises and possible interference of offshore wind farms in the German North Sea. *Marine Ecology Progress Series*. 383, 295–307.

Gimenez, O., Bonner, S.J., King, R., Parker, R.A., Brooks, S.P., Jamieson, L.E., Grosbois, V., Morgan, B.J.T. & Thomas, L. (2009) 'WinBUGS for population ecologists: Bayesian modeling using Markov chain Monte Carlo methods' in David L. Thomson et al. (eds.) Modeling demographic processes in marked populations - Environmental and ecological statistics 3.

Gingerich, P.D. (2015) 'Evolution of whales from land to sea', in Kenneth P. Dial, Neil Shubin, & Elizabeth L. Brainerd (eds.) *Great transformations in verterbrae evolution*. The university of Chicago press. pp. 239–256.

Glennie, R., Buckland, S.T. & Thomas, L. (2015) The effect of animal movement on line transect estimates of abundance. *PLoS ONE,* 10(3), e0121333.

Gonçalves, M.I.C., Sousa-Lima, R.S., Teixeira, N.N., Morete, M.E., Carvalho, G.H., Ferreira, H.M. & Baumgarten, J.E. (2018) Low latitude habitat use patterns of a recovering population of humpback whales. *Journal of the Marine Biological Association of the United Kingdom*. 98 (special issue 5), 1–10.

Götz, T. & Janik, V.M. (2013) Acoustic deterrent devices to prevent pinniped depredation: Efficiency, conservation concerns and possible solutions. *Marine Ecology Progress Series*. 492, 285–302.

Gregr, E.J. & Trites, A.W. (2001) Predictions of critical habitat for five whale species in the waters of coastal British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences*. 58 (7), 1265–1285.

Hair Jr., J.F., Black, W.C., Babin, B.J. & Anderson, R.E. (2014) *Multivariate Data Analysis*. 7th edition. Essex: Pearson.

Halpern, B.S., Walbridge, S., Selkoe, K.A., Kappel, C. V., Micheli, F., D'Agrosa, C., Bruno, J.F., Casey, K.S., Ebert, C., Fox, H.E., Fujita, R., Heinemann, D., Lenihan, H.S., Madin, E.M.P., Perry, M.T., Selig, E.R., Spalding, M., Steneck, R. & Watson, R. (2008) A global map of human impact on Marine Ecosystems. *Science*. 319, 948–952.

Hamazaki, T. (2002) Spatiotemporal prediction models of cetacean habitats in the mid-western North Atlantic Ocean (from Cape Hatteras, North Carolina, U.S.A. to Nova Scotia, Canada). *Marine Mammal Science*. 18 (4), 920–939.

Hammond, P.S. (2010) 'Estimating the abundance of marine mammals', in I. L. Boyd, W. D. Bowen, & S. J. Iverson (eds.) *Marine mammal ecology and conservation: A handbook of techniques*. Oxford: Oxford University Press. pp. 42–67.

Handoh, I.C. & Kawai, T. (2014) Modelling exposure of oceanic higher trophic-level consumers to polychlorinated biphenyls: pollution 'hotspots' in relation to mass mortality events of marine mammals. *Marine Pollution Bulletin*. 85 (2), 824–830.

Hardin, J.W. & Hilbe, J.M. (2012) *Generalized estimating equations*. 2nd edition. Chapman & Hall/CRC.

Harville, D.A. (1977) Maximum Likelihood approaches to variance component estimation and to related problems. *Journal of American Statistical Association*. 72 (358), 320–338.

Hastie, G.D., Swift, R.J., Slesser, G., Thompson, P.M. & Turrell, W.R. (2005) Environmental models for predicting oceanic dolphin habitat in the Northeast Atlantic. *ICES Journal of Marine Science*. 62 (4), 760–770.

Hastie, T.J. & Tibshirani, R. (1990) *Generalized additive models. Monographs on Statistics and Applied Probability* 43. Chapman & Hall/CRC.

Heide-Jørgensen, M.P., Nordøy, E.S., Øien, N., Folkow, L.P., Kleivane, L., Blix, A.S., Jensen, M. V & Laidre, K.L. (2001) Satellite tracking of minke whales (*Balaenoptera acutorostrata*) off the coast of northern Norway. *Journal of Cetacean Research and Management*. 3 (2), 175–178.

Heithaus, M.R. & Dill, L.M. (2006) Does tiger shark predation risk influence foraging habitat use by bottlenose dolphins at multiple spatial scales. *Oikos*. 114 (2), 257–264.

Helm, R., Costa, D., DeBruyn, T., O'Shea, T., Wells, R. & Williams, T. (2015) 'Overview of effects of oil spills on marine mammals', in M Fingas (ed.) *Handbook of oil spill science and technology*. Hoboken, NJ: John Wiley & Sons. pp. 455–475.

Herman, L.M. (2017) The multiple functions of male song within the humpback whale (*Megaptera novaeangliae*) mating system: review, evaluation, and synthesis. *Biological Reviews*. 92, 1795–1818.

Hijmans, R.J. (2016) *Raster: geographic data analysis and modeling. R package version 2.5-8.*

Hofman, R.J. & Bonner, W.N. (1985) Conservation and protection of marine mammals: Past, present and Future. *Marine Mammal Science*. 1 (2), 109–127.

Højsgaard, S., Halekoh, U. & Yan, J. (2016) *Package 'geepack', Version 1.2-1*.

Hoyt, E. (2012) *Marine protected areas for whales dolphins and porpoises: a world handbook for cetacean habitat conservation*. Routledge.

IBAMA (2011) *SisCom: sistema compartilhado de informações ambientais. siscom.ibama.gov.br/ (accessed in 12 Jul 17).*

ICMBio (2017) *PARNA Marinho dos Abrolhos.* Available from: *www.icmbio.gov.br/portal/component/content/article?id=2267:parna-marinho-dos-abrolhos (accessed 7 Feb 2017).*

INEMA-BAHIA (2017) *APA Ponta da Baleia/Abrolhos.* Available from: *www.inema.ba.gov.br/gestao-2/unidades-de-conservacao/apa/apa-ponta-da-baleia-abrolhos/ (accessed 29 Mar 2017).*

Iversen, P.E., Stokke, R., Bloor, P., Unger, S., Nilsen, H.-G. & Jarrah, J. (2009) *Assessment of impacts of offshore oil and gas activities in the North-East Atlantic. Report of the OSPAR Commission, Offshore Industry Service.*

IWC (1998) *Report of the Scientific Committee. Report of the International Whaling Commission 48*. pp. 53–118

IWC (2003) Report of the sub-committee on the comprehensive assessment of humpback whales, report of the scientific committee, annex H. *Journal of Cetacean Research and Management*. 5 (supplement).

IWC (2002) Report of the sub-committee on the comprehensive assessment of North Atlantic humpback whales, report of the scientific committee, annex H. *Journal of Cetacean Research and Management*. 4 (supplement).

IWC (2011) Report of the Workshop on the Comprehensive Assessment of Southern Hemisphere humpback whales, Hobart, Tasmania, 4–7 April 2006. J. *Journal of Cetacean Research and Management*. Special issue 3, 1–50.

Jackson, J.A., Zerbini, A., Clapham, P., Garrigue, C., Hauser, N., Poole, M. & Baker, C.S. (2006) A Bayesian assessment of humpback whales on breeding grounds of eastern Australia and Oceania (IWC Stocks E, E1, E2 and F). SC/A06/HW52. *IWC Workshop on Comprehensive Assessment of Southern Hemisphere Humpback Whales*. 19pp.

Jefferson, T.A., Stacey, P.J. & Baird, R.W. (1991) A review of killer whale interactions with other marine mammals: predation to co-existence. *Mammal Review*. 21 (4), 151–180.

Johnson, D.S., Hooten, M.B. & Kuhn, C.E. (2013) Estimating animal resource selection from telemetry data using point process models. *Journal of animal ecology*. 82., 1155–1164.

Johnston, S.J., Zerbini, A.N. & Butterworth, D.S. (2011) A Bayesian approach to assess the status of Southern hemisphere humpback whales (*Megaptera novaeangliae*) with an application to Breeding Stock G. *Journal of Cetacean Research and Management*. Special issue 3, 309–317.

Jones, E.L., Rendell, L., Pirotta, E. & Long, J.A. (2016) Novel application of a quantitative spatial comparison tool to species distribution data. *Ecological Indicators*. 70, 67–76.

Jones, E.L., Sparling, C.E., McConnell, B.J., Morris, C.D. & Smout, S. (2017) Fine-scale harbour seal usage for informed marine spatial planning. *Scientific Reports*. 7 (11581), 1–11.

Jonsen, I. (2016) Joint estimation over multiple individuals improves behavioural state inference from animal movement data. Scientific Reports. 6 (20625), 1–9.

Jonsen, I.D., Flemming, J.M. & Myers, R.A. (2005) Robust state-space modelling of animal movement data. *Ecology*. 86 (11), 2874–2880.

Jonsen, I.D., McMahon, C.R., Patterson, T.A., Auger-Méthé, M., Harcourt, R., Hindell, M.A. & Bestley, S. (2019) Movement responses to environment: fast inference of variation among southern elephant seals with a mixed effects model. *Ecological Society of America*. 100 (1), e02566.

JPL MUR MEaSUREs Project (2010) *GHRSST level 4 MUR global foundation sea surface temperature analysis, PO.DAAC, Ver. 2.* Available from: *podaac.jpl.nasa.gov/dataset/JPL-L4UHfnd-GLOB-MUR (accessed in 14 Jun 2016)*.

Kahle, D. & Wickham, H. (2013) ggmap: spatial visualization with ggplot2. *The R Journal*. 5 (1), 144–161.

Kajiwara, N., Matsuoka, S., Iwata, H., Tanabe, S., Rosas, F.C.W., Fillmann, G. & Readman, J.W. (2004) Contamination by Persistent Organochlorines in Cetaceans Incidentally Caught Along Brazilian Coastal Waters. *Archives of Environmental Contamination and Toxicology*. 46 (1), 124–134.

Kanaji, Y., Okazaki, M. & Miyashita, T. (2017) Deep-Sea Research II - Spatial patterns of distribution, abundance, and species diversity of small odontocetes estimated using density surface modeling with line transect sampling. *Deep-Sea Research Part II*. 140, 151–162.

Keating, K.A. & Cherry, S. (2004) Use and interpretation of logistic regression in habitat-selection studies. *Journal of Wildlife Management*. 68 (4), 774–789.

Keller, C.A., Ward-Geiger, L.I., Brooks, W.B., Slay, C.K., Taylor, C.R. & Zoodsma, B.J. (2006) North Atlantic right whale distribution in relation to sea-surface temperature in the southeastern United States calving grounds. *Marine Mammal Science*. 22 (2), 426–445.

Kileso, A., Chubarenko, B., Zemlys, P. & Kuzmenko, I. (2014) Oil spill modelling methods: application to the south–eastern part of the Baltic Sea. *Baltica*. 27, 15–22.

King, S.L., Schick, R.S., Donovan, C., Booth, C.G., Burgman, M., Thomas, L. & Harwood, J. (2015) An interim framework for assessing the population consequences of disturbance. *Methods in Ecology and Evolution*. 6, 1150–1158.

Kite-powell, H.L., Knowlton, A. & Brown, M. (2007) *Modeling the effect of vessel speed on Right Whale ship strike risk*. Report for NOAA/NMFS, Project NA04NMF47202394.

Kleindienst, S., Paul, J.H. & Joye, S.B. (2015) Using dispersants after oil spills: impacts on the composition and activity of microbial communities. *Nature Reviews Microbiology*. doi:10.1038/nrmicro3452.

Knowlton, A.R. & Kraus, S.D. (2001) Mortality and serious injury of northern right whales (*Eubalaena glacialis*) in the western North Atlantic Ocean. *Journal of Cetacean Research and Management*. Special issue 2, 193–208.

Kruschke, J. (2014) *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan* . 2nd edition. Academic Press, Elsevier.

Laake, J.L., Punt, A.E., Hobbs, R., Ferguson, M., Rugh, D. & Breiwick, J. (2012) Gray whale southbound migration surveys 1967-2006: an integrated re-analysis. *Journal of Cetacean Research and Management*. 12 (3), 287–306.

Laist, D.W. (1997) 'Impacts of marine debris: Entanglement of marine life in marine debris including a comprehensive list of species with entanglement and ingestion records', in James M. Coe & Donald B. Rogers (eds.) *Marine Debris: Sources, Impacts and Solutions*. New York: Springer. pp. 99–139.

Laist, D.W., Knowlton, A.R., Mead, J.G., Collet, A.S. & Podestà, M. (2001) Collisions between ships and whales. *Marine Mammal Science*. 17 (1), 35–75.

Large, W.G., & Yeager, S.G. (2009) The global climatology of an interannually varying air-sea flux data set. *Climate Dynamics*. 33 (2–3), 341–364.

Lima, S.L. (1998) Stress and decision-making under the risk of predation: recent developments from behavioral, reproductive and ecological perspectives. *Advances in the Study of Behaviour*. 27, 215–290.

Lindsay, R.E., Constantine, R., Robbins, J., Mattila, D.K., Tagarino, A. & Dennis, T.E. (2016) Characterising essential breeding habitat for whales informs the development of large-scale Marine Protected Areas in the South Pacific. *Marine Ecology Progress Series*. 548, 263–275.

Lodi, L. (1994) Ocorrências de baleias-jubarte no Arquipélago de Fernando de Noronha, incluindo um resumo de registros de capturas no Nordeste do Brasil. *Biotemas*. 7 (1 and 2), 123–124.

López, B.D., Dolphin, B., Beiramar, A. & Grove, O. (2018) 'Hot deals at sea': responses of a top predator (Bottlenose dolphin, *Tursiops truncatus*) to human-induced changes in the coastal ecosystem. *Behavioral Ecology*. *Early view*, 1–10.

Lucena, M., Fluminense, U.F., Sissini, M. & Sazima, I. (2016) Out of the mainstream: humpback whale calving site and associated fishes at an oceanic island off Brazil. *Marine Biodiversity*. 46 (27), 1–3.

Macleod, K., Fairbairns, R., Gill, A., Fairbairns, B., Gordon, J., Blair-Myers, C. & Parsons, E.C.M. (2004) Seasonal distribution of minke whales Balaenoptera acutorostrata in relation to physiography and prey off the Isle of Mull, Scotland. *Marine Ecology Progress Series*. 277, 263–274.

Manabe, S. (1969) Climate and the ocean circulation: 1. the atmospheric circulation and the effect of heat transfer by ocean currents. *Monthly Weather Review*. 97 (11), 739–774.

Manly, B.F.J., McDonald, L.L., Thomas, D.L., McDonald, T.L. & Erickson, W. (2002) *Resource selection by animals: statistical design and analysis for field studies*. 2nd edition. New York: Chapman & Hall.

Mannocci, L., Monestiez, P., Spitz, J. & Ridoux, V. (2015) Extrapolating cetacean densities beyond surveyed regions: habitat-based predictions in the circumtropical belt. *Journal of Biogeography*. 42, 1267–1280.

Marchioro, G.B., Nunes, M.A., Dutra, G.F., Moura, R.L. & Pereira, P.G.P. (2005) Avaliação dos impactos da exploração e produção de hidrocarbonetos no Banco dos Abrolhos e adjacências. *Megadiversidade*. 1 (2), 225–310.

Marsh, H. & Sinclair, D.F. (1989) Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. *The Journal of Wildlife Management*. 53 (4), 1017–1024.

Marshall, C.R. & Quental, T.B. (2016) The uncertain role of diversity dependence in species diversification and the need to incorporate time-varying carrying capacities. *Philosophical Transactions of the Royal Society B*. 371, 20150217.

Martins, C.C.A., Andriolo, A., Engel, M.H., Kinas, P.G. & Saito, C.H. (2013) Identifying priority areas for humpback whale conservation at Eastern Brazilian Coast. *Ocean and Coastal Management*. 75, 63–71.

Martins, C.C.A., Morete, M.E., Engel, M.H., Freitas, A.C., Secchi, E.R. & Kinas, P.G. (2001) Aspects of habitat use patterns of humpback whales in the Abrolhos Bank, Brazil, breeding ground. *Memoirs of the Queensland Museum*. 47 (2), 563–570.

Matkin, C.O., Saulitis, E.L., Ellis, G.M., Olesiuk, P. & Rice, S.D. (2008) Ongoing population-level impacts on killer whales *Orcinus orca* following the 'Exxon Valdez' oil spill in Prince William Sound, Alaska. *Marine Ecology Progress Series*. 356, 269–281.

Matthiopoulos, J. (2003) The use of space by animals as a function of accessibility and preference. *Ecological Modelling*. 159, 239–268.

Matthiopoulos, J. & Aarts, G. (2010) 'The spatial analysis of marine mammal abundance', in Ian L. Boyd, W. Don Bowen, & Sara J. Iverson (eds.) *Marine Mammal Ecology and Conservation: A handbook of techniques*. London: Oxford University Press. pp. 68–97.

Mcallister, M.K., Pikitch, E.K., Punt, A.E. & Hilborn, R. (1994) A Bayesian approach to stock assessment and harvest decisions using the sampling/importance resampling algorithm. *Canadian Journal of Fisheries and Aquatic Sciences*. 51, 2673–2687.

McClintock, B.T., London, J.M., Cameron, M.F. & Boveng, P.L. (2015) Modelling animal movement using the Argos satellite telemetry location error ellipse. *Methods in Ecology and Evolution.* 6, 266–277.

McKenna, M.F., Katz, S.L., Condit, C. & Walbridge, S. (2012) Response of commercial ships to a voluntary speed reduction measure: are voluntary strategies adequate for mitigating ship-strike risk? *Coastal Management*. 40 (6), 634–650.

McMillan, M., Shepherd, A., Sundal, A., Briggs, K., Muir, A., Ridout, A., Hogg, A. & Wingham, D. (2015) Increased ice losses from Antarctica detected by CryoSat-2. *Geophysical Research Letters*. 41, 3899–3905.

Mei, Z., Chen, M., Li, Y., Huang, S., Haung, J., Han, Y., Zhu, B., Li, C., Wang, K. & Wang, D. (2017) Habitat preference of the Yangtze finless porpoise in a minimally disturbed environment. *Ecological Modelling*. 353, 47–53.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N. & Teller, A.H. (1953) Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*. 21, 1087–1092.

Meÿer, M.A., Best, P.B., Anderson-Reade, M.D., Cliff, G., Dudley, S.F.J. & Kirkman, S.P. (2011) Trends and interventions in large whale entanglement along the South African coast. *African Journal of Marine Science*. 33 (3), 429–439.

Miller, D.L. (2016) *Distance: distance sampling detection function and abundance estimation. R package version 0.9.6.*

Miller, D.L., Burt, M.L., Rexstad, E.A. & Thomas, L. (2013) Spatial models for distance sampling data: recent developments and future directions. *Methods in Ecology and Evolution*. 4, 1001–1010.

Miller, D.L., Rexstad, E., Burt, L., Bravington, M. V. & Hedley, S.L. (2017) *dsm: density surface modelling of distance sampling data. R package version 2.2.14.*

Molinari, R.L., Olson, D. & Reverdin, G. (1990) Surface current distributions in the tropical Indian Ocean derived from compilations of surface buoy trajectories. *Journal of Geophysical Research*. 95 (C5), 7217–7238

Morais, I.O.B., Danilewicz, D., Zerbini, A.N., Edmundson, W., Hart, I.B. & Bortolotto, G.A. (2017) From the southern right whale hunting decline to the humpback whaling expansion: a review of whale catch records in the tropical western South Atlantic Ocean. *Mammal Review*. 47, 11–23.

Morete, M.E., Pace, R.M., Martins, C.C.A., Freitas, A.C. & Engel, M.H. (2003) Indexing seasonal abundance of humpback whales around Abrolhos Archipelago, Bahia, Brazil. *Latin American Journal of Aquatic Mammals*. 2 (1), 21–28.

Morissette, L., Hammil, M.O. & Savenkoff, C. (2006) The trophic role of marine mammals in the northern gulf of St. Lawrence. *Marine Mammal Science*. 22 (1), 74–103.

Moura, J.F., Rodrigues, D.P., Roges, E.M., Souza, R.L., Ott, P.H., Tavares, M., Lemos, L.S., Tavares, D.C. & Siciliano, S. (2013) Humpback whales washed ashore in southeastern Brazil from 1981 to 2011: stranding patterns and microbial pathogens survey. *Biologia*. 68 (5), 992–999.

New, L.F., Clark, J.S., Costa, D.P., Fleishman, E., Hindell, M.A., Klanjšček, T., Lusseau, D., Kraus, S., McMahon, C.R., Robinson, P.W., Schick, R.S., Schwarz, L.K., Simmons, S.E., Thomas, L., Tyack, P. & Harwood, J. (2014) Using short-term measures of behaviour to estimate long-term fitness of southern elephant seals. *Marine Ecology Progress Series*. 496, 99–108.

Newman, K., Buckland, S.T., Morgan, B., King, R., Borchers, D.L., Cole, D., Besbeas, P., Gimenez, O. & Thomas, L. (2014) *Modelling Population Dynamics: Model Formulation, Fitting and Assessment Using State-Space Methods*. Springer.

Nichols, O.C. & Kite-Powell, H.L. (2005) *Analysis of risk to north Atlantic right whales* (Eubalaena glacialis) *from shipping traffic in Cape Cod Bay, final report submitted to NOAA fisheries,* NA03NMF4720489.

Nicol, S. (2006) Krill, currents, and sea ice: *Euphausia superba* and its changing environment. *BioScience*. 56 (2), 111–120.

Nowacek, D.P., Thorne, L.H., Johnston, D.W. & Tyack, P.L. (2007) Responses of cetaceans to anthropogenic noise. *Mammal Review*. 37 (2), 81–115.

Ott, P.H., Milmann, L., Santos, M.C.O., Rogers, E.M., Rodrigues, D.P. & Siciliano, S. (2016) *Humpback whale breeding stock 'A': increasing threats to a recently down-listed species off Brazilian fauna. Work SC/66b/SH/04 presented to the International Whaling Commission. Scientific Committee, Bled, Slovenia*.

Oviedo, L. & Solís, M. (2008) Underwater topography determines critical breeding habitat for humpback whales near Osa Peninsula, Costa Rica: implications for Marine Protected Areas. *Revista de Biología Tropical*. 56, 591–602.

Pan, W. (2001) Akaike's information criterion in generalized estimating equations. *Biometrics*. 57 (1), 120–125.

Panigada, S., Zanardelli, M., Canese, S. & Jahoda, M. (1999) How deep can baleen whales dive? *Marine Ecology Progress Series*. 187, 309–311.

Pardo, M.A., Gerrodette, T., Beier, E., Gendron, D., Forney, K.A., Chivers, S.J., Barlow, J. & Palacios, D.M. (2015) Inferring cetacean population densities from the absolute dynamic topography of the ocean in a hierarchical Bayesian framework. *Plos One*. 10 (3), e0120727.

Wade, P.R. (2018) 'Stock assessment', in Bernd Würsig, J.G.M. Thewissen, & Kit Kovacs (eds.) *Encyclopedia of Marine Mammals*. 3rd edition. London: Academic Press, Elsevier. pp. 938–941.

Pauly, D., Christensen, V., Dalsgaard, J., Froese, R. & Torres Jr, F. (1998) Fishing down marine food webs. *Science*. 279, 860–863.

Pavanato, H.J., Mayer, F.P., Wedekin, L.L., Engel, M.H. & Kinas, P.G. (2018) Prediction of humpback whale group densities along the Brazilian coast using spatial autoregressive models. *Marine Mammal Science*. 34 (3), 1–21.

Pavanato, H.J., Wedekin, L.L., Guilherme-Silveira, F.R., Engel, M.H. & Kinas, P.G. (2017) Estimating humpback whale abundance using hierarchical distance sampling. *Ecological Modelling*. 358, 10–18.

Pella, J.J. & Tomlinson, P.K. (1969) A generalised stock production model. *Inter-American Tropical Tuna Commission Bulletin*. 13 (3), 421–496.

Peltier, H., Jepson, P.D., Dabin, W., Deaville, R., Daniel, P., Van Canneyt, O. & Ridoux, V. (2014) The contribution of stranding data to monitoring and conservation strategies for cetaceans: Developing spatially explicit mortality indicators for common dolphins (*Delphinus delphis*) in the eastern North-Atlantic. *Ecological Indicators*. 39, 203–214.

Pen, O., Thorne, L.H., Foley, H.J., Baird, R.W., Webster, D.L., Swaim, Z.T. & Read, A.J. (2017) Movement and foraging behavior of short-finned pilot whales in the Mid-Atlantic Bight: importance of bathymetric features and implications for management. *Marine Ecology Progress Series*. 584, 245–257.

Pirotta, E., Booth, C.G., Costa, D.P., Fleishman, E., Kraus, S.D., Lusseau, D., Moretti, D., New, L.F., Schick, R.S., Schwarz, L.K., Simmons, S.E., Thomas, L., Tyack, P.L., Weise, M.J., Wells, R.S. & Harwood, J. (2018) Understanding the population consequences of disturbance. *Ecology and Evolution*. 8, 9934–9946.

Pirotta, E., Matthiopoulos, J., MacKenzie, M., Scott-Hayward, L. & Rendell, L. (2011) Modelling sperm whale habitat preference: A novel approach combining transect and follow data. *Marine Ecology Progress Series*. 436, 257–272.

QGIS development team (2015) *QGIS Geographic Information System. Open Source Geospatial Foundation Project: http://www.qgis.org/*.

R Core Team (2015) *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

R Core Team (2017) *R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/*.

Rasmussen, K., Palacios, D.M., Rosa, L.D., Secchi, E.R., Steiger, G.H., Allen, J.M. & Stone, G.S. (2007) Southern Hemisphere humpback whales wintering off Central America: insights from water temperature into the longest mammalian migration. *Biology Letters*. 3, 302–305.

Rayment, W., Dawson, S. & Webster, T. (2015) Breeding status affects fine-scale habitat selection of southern right whales on their wintering grounds. *Journal of Biogeography*. 42 (3), 463–474.

Read, A.J., Drinker, P. & Northridge, S. (2006) Bycatch of marine mammals in U.S. and global fisheries. *Conservation Biology*. 20 (1), 163–169.

Redfern, J.V., Ferguson, M.C., Becker, E.A., Hyrenbach, K.D., Good, C., Barlow, J., Kaschner, K., Baumgartner, M.F., Forney, K.A., Ballance, L.T., Fauchald, P., Halpin, P., Hamazaki, T., Pershing, A.J., Qian, S.S., Read, A., Reilly, S.B., Torres, L. & Werner, F. (2006) Techniques for cetacean-habitat modeling. *Marine Ecology Progress Series*. 310, 271–295.

Reisinger, R.R., Oosthuizen, W.C., Péron, G., Toussaint, D.C., Andrews, R.D. & De Bruyn, P.J.N. (2014) Satellite tagging and biopsy sampling of killer whales at subantarctic Marion Island: effectiveness, immediate reactions and long-term responses. *PLoS ONE*. 9 (11), 1–14.

Renner, I.W., Elith, J., Baddeley, A., Fithian, W., Hastie, T., Phillips, S.J., Popovic, G. & Warton, D.I. (2015) Point process models for presence-only analysis. *Methods in Ecology and Evolution*. 6 (4), 366–379.

Richardson, W.J. & Würsig, B. (1997) Influences of man-made noise and other human actions on cetacean behaviour. *Marine and Freshwater Behaviour and Physiology*. 29 (1–4), 183–209.

Robbins, J. & Mattila, D.K. (2004) *Estimating humpback whale entanglement rates on the basis of scar evidence. Report to the Northeast Fisheries Science Center*.

Roberts, J.J., Best, B.D., Mannocci, L., Fujioka, E., Halpin, P.N., Palka, D.L., Garrison, L.P., Mullin, K.D., Cole, T.V.N., Khan, C.B., McLellan, W.A., Pabst, D.A. & Lockhart, G.G. (2016) Habitat-based cetacean density models for the U.S. Atlantic and Gulf of Mexico. *Scientific Reports*. 6, 22615.

Rocha-Campos, C.C., Moreno, I.B., Rocha, J.M., Palazzo Jr, J.T., Groch, K.R., Oliveira, L.R., Gonçalves, L., Engel, M.H., Marcondes, M.C.C., Muelbert, M.M.C., Ott, P.H. & Silva, V.M.F. (2011) *Plano de ação nacional para conservação dos mamíferos aquáticos: grandes cetáceos e pinípedes. Série Espécies Ameaçadas nº 14*. Claudia Cavalcante Rocha-Campos & Ibsen de Gusmão Câmara (eds.). Brasília: Instituto Chico Mendes de Conservação da Biodiversidade, ICMBio.

Rocha, R.C., Clapham, P.J. & Ivashchenko, Y.V. (2014) Emptying the oceans: a summary of industrial whaling catches in the 20th century. *Marine Fisheries Review*. 76 (4), 37–48.

Roman, J., Estes, J.A., Morissette, L., Smith, C., Costa, D., Mccarthy, J., Nation, J.B., Nicol, S., Pershing, A. & Smetacek, V. (2014) Whales as marine ecosystem engineers. *Frontiers in Ecology and the Environment*. 12 (7), 377–385.

Romero, A.F., Abessa, D.M.S., Fontes, R.F.C. & Silva, G.H. (2013) Integrated assessment for establishing an oil environmental vulnerability map: case study for the Santos Basin region, Brazil. *Marine Pollution Bulletin*. 74 (1), 156–164.

Roncon, G., Bestley, S., Mcmahon, C.R., Wienecke, B. & Hindell, M.A. (2018) View from below: inferring behavior and physiology of Southern Ocean marine predators from dive telemetry. *Frontiers in Marine Science*. 5 (464), 1–23.

Ronconi, R.A., Allard, K.A. & Taylor, P.D. (2015) Bird interactions with offshore oil and gas platforms: review of impacts and monitoring techniques. *Journal of Environmental Management*. 147, 34–45.

Rosenbaum, H.C., Pomilla, C.C., Leslie, M.C., Mendez, M.C., Best, P.B., Collins, T., Engel, M.H., Ersts, P.J., Findlay, K.P., Bonatto, S., Kotze, P.G.H., Meyer, M., Minton, G., Barendse, J., Thorton, M., Razafindrakoto, Y. & Ngouessono, S. (2006) Mitochondrial DNA diversity and population structure of humpback whales from their wintering areas (breeding stocks) in the Indian and South Atlantic Oceans (wintering regions A, B, C and X). SC/A06/HW41. *IWC Workshop on Comprehensive Assessment of Southern Hemisphere Humpback Whales*. 16pp.

Rossi-Santos, M.R. (2015) Oil industry and noise pollution in the humpback whale (*Megaptera novaeangliae*) soundscape ecology of the Southwestern Atlantic breeding ground. *Journal of Coastal Research*. 31 (1), 184–195.

Rossi-Santos, M.R., Neto, E.S., Baracho, C.G., Cipolotti, S.R., Marcovaldi, E. & Engel, M.H. (2008) Occurrence and distribution of humpback whales (*Megaptera novaeangliae*) on the north coast of the State of Bahia, Brazil, 2000-2006. *ICES Journal of Marine Science*. 65 (4), 667–673.

Runge, C.A., Watson, J.E.M., Butchart, S.H.M., Hanson, J.O., Possingham, H.P. & Fuller, R.A. (2015) Protected areas and global conservation of migratory birds. *Science*. 350 (6265), 1255–1258.

Ruoppolo, V., Vanstreels, R.E.T., Silveira, L.F., Zerbini, A.N., Colman, L., Woehler, E.J., Nascimento, C.C., Silva-Filho, R.P., Pina, M.C., Martins, P. & Ranieri, A. (2017) Sensitivity mapping for oil spill response: a comprehensive framework to identify wildlife and areas at risk along the coastline of Brazil. *International Oil Spill Conference Proceedings*2017.

Russell, D.J.F., Hastie, G.D., Thompson, D., Janik, V., Hammond, P.S., Scott-Haywood, L., Matthiopoulos, J., Jones, E. & McConnell, B.J. (2016) Avoidance of windfarms by harbour seals is limited to pile driving activities. *Journal of Applied Ecology*. doi: 10.1111/1365-2664.12678.

Schaffar, A., Madon, B., Garrigue, C. & Constantine, R. (2013) Behavioural effects of whale-watching activities on an Endangered population of humpback whales wintering in New Caledonia. *Endangered Species Research*. 19, 245–254.

Scheidat, M., Tougaard, J., Brasseur, S., Carstensen, J., Van Polanen Petel, T., Teilmann, J. & Reijnders, P. (2011) Harbour porpoises (*Phocoena phocoena*) and wind farms: a case study in the Dutch North Sea. *Environmental Research Letters*. 6 (2), 025102.

Scott-Hayward, L., Oedekoven, C., Mackenzie, M., Walker, C. & Rexstad, E. (2017) *MRSea Package (version 1.0-beta): statistical modelling of bird and cetacean distributions in offshore renewables development areas. University of St. Andrews: Contract with Marine Scotland: SB9 (CR/2012/05)*.

Secchi, E.R., Ott, P.H. & Danilewicz, D. (2002) Report of the fourth workshop for the coordinated research and conservation of the franciscana dolphin (*Pontoporia blainvillei*) in the western South Atlantic. *Latin American Journal of Aquatic Mammals*. 1 (1), 11–20.

Shelden, K.E.W., Rugh, D.J. & Schulman-Janiger, A. (2004) Gray whales born north of Mexico: indicator of recovery or consequence of regime shift? *Ecological Applications*. 14 (6), 1789–1805.

Siciliano, S. (1997) *Características da população de baleias-jubarte (*Megaptera novaeangliae*) da costa Brasileira, com especial referência aos Bancos de Abrolhos. Masters dissertation, Universidade Federal Rural do Rio de Janeiro, Rio de Janeiro, RJ*.

Siciliano, S., Moura, J.F., Filgueiras, H.R., Rodrigues, P.P. & Oliveira Leite Jr., N. (2012) Sightings of humpback whales on the Vitória-Trindade chain and around Trindade Island, Brazil. *Brazilian Journal of Oceanography*. 60 (3), 455–459.

Sing, T., Sander, O., Beerenwinkel, N. & Lengauer, T. (2009) *ROCR: visualizing classifier performance in R. R package, version 1.0-4.*

Slijper, E.J. (1962) *Whales. Hutchinson, London.* London: Hutchinson.

Smith, T.D., Allen, J., Clapham, P.J., Hammond, P.S., Katona, S., Larsen, F., Lien, J., Mattila, D., Palsbøll, P.J., Sigurjónsson, J., Stevick, P.T. & Øien, N. (1999) An ocean-basin-wide mark-recapture study of the North Atlantic humpback whale (*Megaptera novaeangliae*). *Marine Mammal Science*. 15 (1), 1–32.

Smultea, M.A. (1994) Segregation by humpback whale (*Megaptera novaeangliae*) cows with a calf in coastal habitat near the island of Hawaii. *Canadian Journal of Zoology*. 72 (5), 805–811.

Smultea, M.A. & Würsig, B. (1995) Behavioral reactions of bottlenose dolphins to the Mega Borg oil spill, Gulf of Mexico 1990. *Aquatic Mammals*. 21 (3), 171–181.

Sorensen, P.W., Medved, R.J., Hyman, M.A.M. & Winn, H.E. (1984) Distribution and abundance of cetaceans in the vicinity of human activities along the continental shelf of the northwestern Atlantic. *Marine Environmental Research*. 12, 69–81.

Sousa-Lima, R.S. & Clark, C.W. (2009) Whale sound recording technology as a tool for assessing the effects of boat noise in a Brazilian marine park. *Park Science*. 26 (1), 59–63.

Southall, B.L., Bowles, A.E., Ellison, W.T., Finneran, J.J., Gentry, R.L., Greene Jr., C.R., Kastak, D., Ketten, D.R., Miller, J.H., Nachtigall, P.E., Richardson, W.J., Thomas, J.A. & Tyack, P.L. (2007) Marine mammal noise exposure criteria: Initial Scientific Recommendations. *Aquatic Mammals*. 33 (4), 411–521.

Spaulding, M.L. (2017) State of the art review and future directions in oil spill modeling. *Marine Pollution Bulletin*. 115 (1–2), 7–19.

Stevick, P.T., Allen, J., Clapham, P.J., Friday, N., Katona, S.K., Larsen, F., Lien, J., Mattila, D.K., Palsbøll, P.J., Sigurjónsson, J., Smith, T.D., Øien, N. & Hammond, P.S. (2003) North Atlantic humpback whale abundance and rate of increase four decades after protection from whaling. *Marine Ecology Progress Series*. 258, 263–273.

Stevick, P.T., Berrow, S.D., Bérube, M., Bouveret, L., Broms, F., Jann, B., Kennedy, A., Suárez, P.L., Meunier, M., Ryan, C. & Wenzel, F. (2016) There and back again: multiple and return exchange of humpback whales between breeding habitats separated by an

ocean basin. *Journal of the Marine Biological Association of the United Kingdom*. 96 (4), 885–890.

Stevick, P.T., Godoy, L.P., Mcosker, M., Engel, M.H. & Allen, J. (2006) A note on the movement of a humpback whale from Abrolhos Bank, Brazil to South Georgia. *Journal of Cetacean Research and Management*. 8 (3), 297–300.

Sumner, M.D. (2016) trip: tools for the analysis of animal track data. R package version 1.5.0.

Taber, S. & Thomas, P. (1982) Calf development and mother-calf spatial relationships in southern right whales. *Animal Behaviour*. 30, 1072–1083.

Taylor, B.L., Rojas-Bracho, L., Moore, J., Jaramillo-Legorreta, A., Ver Hoef, J.M., Cardenas-Hinojosa, G., Nieto-Garcia, E., Barlow, J., Gerrodette, T., Tregenza, N., Thomas, L. & Hammond, P.S. (2017) Extinction is imminent for Mexico's endemic porpoise unless fishery bycatch is eliminated. *Conservation letters*. 10 (5), 588–595.

Torres, L.G., Smith, T.D., Sutton, P., MacDiarmid, A., Bannister, J. & Miyashita, T. (2013) From exploitation to conservation: habitat models using whaling data predict distribution patterns and threat exposure of an endangered whale. *Diversity and Distributions*. 19, 1138–1152.

Townsend, C.H. (1935) The distribution of certain whales as shown by logbook records of American whaleships. *Zoologica*. 19 (1), 3–18.

Trudelle, L., Cerchio, S., Zerbini, A.N., Geyer, Y., Mayer, F.-X., Jung, J., Hervé, M.R., Pous, S., Adam, O. & Charrassin, J.-B. (2016) Influence of environmental parameters on movements and habitat utilization of humpback whales (*Megaptera novaeangliae*) in the Madagascar breeding ground. *Royal Society Open Science*. 3, 160616.

Venables, W.N. & Ripley, B.D. (2002) *Modern Applied Statistics with S, 4 edition*. Springer.

Víkingsson, G.A. & Heide-Jorgensen, M.P. (2015) First indications of autumn migration routes and destination of common minke whales tracked by satellite in the North Atlantic during 2001-2011. *Marine Mammal Science*. 31 (1), 376–385.

Virgili, A., Authier, M., Boisseau, O., Cañadas, A., Claridge, D., Cole, T., Corkeron, P., Dorémus, G., David, L., Di, N., Charlotte, M., Tim, D., García, I., Sophie, B., Giancarlo, L., Mark, L., Saavedra, C., Santos, M.B., Van Canneyt, O., et al. (2018) Combining multiple visual surveys to model the habitat of deep-diving cetaceans at the basin scale: large-scale modelling of deep-diving cetacean habitats. *Global Ecology and Biogeography*. 1–15. 10.1111/geb.12850

Wade, P.R. (2002) A Bayesian stock assessment of the eastern Pacific gray whale using abundance and harvest data from 1967-1996. *Journal of cetacean research and management*. 4 (1).

Van Waerebeek, K., Baker, A.N., Félix, F., Gedamke, J., Iñiguez, M., Sanino, G.P., Secchi, E., Sutaria, D., Van Helden, A. & Wang, Y. (2007) Vessel collisions with small cetaceans worldwide and with large whales in the Southern Hemisphere, an initial assessment. *Latin American Journal of Aquatic Mammals*. 6 (1), 43–69.

Walker, C.G., Mackenzie, M.L., Donovan, C.R. & O'Sullivan, M.J. (2011) SALSA: a spatially adaptive local smoothing algorithm. *Journal of Statistical Computation and Simulation*. 81 (2), 179–191.

Wallace, B.P., Brosnan, T., McLamb, D., Rowles, T., Ruder, E., Schroeder, B., Schwacke, L., Stacy, B., Sullivan, L., Takeshita, R. & Wehner, D. (2017) Effects of the Deepwater Horizon oil spill on protected marine species. *Endangered Species Research*. 33 (1), 1–7.

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J.C., Fromentin, J.-M., Hoegh-Guldberg, O. & Bairlein, F. (2002) Ecological responses to recent climate change. *Nature*. 416, 389–395.

Webster, P.J., Holland, G.J., Curry, J.A. & Chang, H. (2005) Changes in tropical cyclone number, duration, and intensity in a warming environment. *Science*. 309, 1844–1846.

Wedekin, L.L. (2011) *Ecologia populacional da baleia-jubarte (Megaptera novaeangliae Borowski, 1871) em sua área reprodutiva na costa do Brasil, Oceano Atlântico Sul. PhD dissertation, Universidade Federal do Paraná, Curitiba, PR*. p.1–166.

Wedekin, L.L., Engel, M.H., Andriolo, A., Prado, P.I., Zerbini, A.N., Marcondes, M.M.C. & Kinas, P.G. (2017) Running fast in the slow lane: rapid population growth of humpback whales after exploitation. *Marine Ecology Progress Series*. 575, 195–206.

Wedekin, L.L., Neves, M.C., Marcondes, M.C.C., Baracho, C., Rossi-Santos, M.R., Engel, M.H. & Simões-Lopes, P.C. (2010) Site fidelity and movements of humpback whales (*Megaptera novaeangliae*) on the Brazilian breeding ground, southwestern Atlantic. *Marine Mammal Science*. 26 (4), 787–802.

Wedekin, L.L., Rossi-Santos, M.R., Baracho, C., Cypriano-Souza, A.L. & Simões-Lopes, P.C. (2014) Cetacean records along a coastal-offshore gradient in the Vitória-Trindade Chain, western South Atlantic Ocean. *Brazilian Journal of Biology*. 74 (1), 137–144.

Wenzel, F.W., Allen, J., Berrow, S., Hazevoet, C.J., Jann, B., Seton, R.E., Steiner, L., Stevick, P., Suárez, P.L. & Whooley, P. (2009) Current knowledge on the distribution and relative abundance of humpback whales (*Megaptera novaeangliae*) off the Cape Verde islands, eastern North Atlantic. *Aquatic Mammals*. 35 (4), 502–510.

Werner, T., Pinto, L.P., Dutra, G.F. & Pereira, P.G.P. (2000) Abrolhos 2000: conserving the Southern Atlantic's richest coastal biodiversity into the next century. *Coastal Management*. 28, 99–108.

Whitehead, H. & Jonsen, I.D. (2013) *Inferring Animal Densities from Tracking Data Using Markov Chains*. 8 (4), 1–5.

Whitehead, H. & Moore, M.J. (1982) Distribution and movements of West Indian humpback whales in winter. Canadian Journal of Zoology 60:2203-2211. *Canadian Journal of Zoology*. 60, 2203–2211.

Wickham, H. (2009) *ggplot2: elegant graphics for data analysis*. New York: Springer-Verlag.

Wickham, H. (2011) The Split-Apply-Combine strategy for data analysis. *Journal of Statistical Software*. 40 (1), 1–29.

Williams, R. & O'Hara, P. (2010) Modelling ship strike risk to fin, humpback and killer whales in British Columbia, Canada. *Journal of Cetacean Research and Management*. 11 (1), 1–8.

Williamson, G.R. (1975) Minke whales off Brazil. *Scientific Report of the Whales Research Institute*. 27, 37–59.

Wilson, B., Thompson, P.M. & Hammond, P.S. (1997) Habitat use by bottlenose dolphins: seasonal distribution and stratified movement patterns in the Moray Firth, Scotland. *Journal of Applied Ecology*. 34 (6), 1365–1374.

Wirsing, A.J., Heithaus, M.R., Frid, A. & Dill, L.M. (2008) Seascapes of fear: Evaluating sublethal predator effects experienced and generated by marine mammals. *Marine Mammal Science*. 24 (1), 1–15.

Wolfson, A., VanBlaricom, G.R., Davis, N. & Lewbel, G.S. (1979) The Marine Life of an Offshore Oil Platform. *Marine Ecology Progress Series*. 1, 181–89.

Wood, S.N. (2006) *Generalized additive models: an introduction with R*. Boca Raton, Florida: Chapman & Hall/CRC.

Wood, S.N. (2017) *Generalized additive models: an introduction with R*. 2nd edition. Chapman & Hall/CRC.

Zemsky, V.A., Berzin, A.A., Mikhalev, Y.A. & Tormosov, D.D. (1996) *Soviet Antarctic whaling data (1947–1972). 2nd ed.*

Zerbini, A.N. (2005) An updated Bayesian assessment of the Southern Hemisphere humpback whale Breeding Stock A. SC/57/SH16. *Report of the IWC Scientific Committee* 13.

Zerbini, A.N., Andriolo, A., Heide-Jørgensen, M.P., Pizzorno, J.L., Maia, Y.G., VanBlaricom, G.R., Demaster, D.P., Simões-Lopes, P.C., Moreira, S. & Bethlem, C.

(2006) Satellite-monitored movements of humpback whales *Megaptera novaeangliae* in the Southwest Atlantic Ocean. *Marine Ecology Progress Series*. 313, 295–304.

Zerbini, A.N., Andriolo, A., Moreira, S., Maia, Y.G., VanBlaricom, G.R. & Demaster, D.P. (2006) 'Movimentos e utilização do habitat de baleias jubarte (*Megaptera novaeangliae*) monitoradas por satélite no litoral do Brasil e suas implicalções para a indústria do petróleo e gás-natural', in *Anais da Rio Oil & Gas Expo and Conference 2006*. 2006. pp. 1–8.

Zerbini, A.N., Andriolo, A., Rocha, J.M., Simões-Lopes, P.C., Siciliano, S., Pizzorno, J.L., Waite, J.M., Demaster, D.P. & VanBlaricom, G.R. (2004) Winter distribution and abundance of humpback whales (*Megaptera novaeangliae*) off Northeastern Brazil. *Journal of Cetacean Research and Management*. 6 (1), 101–107.

Zerbini, A.N., Baumgartner, M., Kennedy, A., Rone, B., Wade, P. & Clapham, P.J. (2015) Space use patterns of the endangered North Pacific right whale *Eubalaena japonica* in the Bering Sea. *Marine Ecology Progress Series*. 532, 269–281.

Zerbini, A.N., Clapham, P.J. & Wade, P.R. (2010) Assessing plausible rates of population growth in humpback whales from life-history data. *Marine Biology*. 157 (6), 1225–1236.

Zerbini, A.N., Ward, E.J., Kinas, P.G., Engel, M.H. & Andriolo, A. (2011) A Bayesian assessment of the conservation status of humpback whales (*Megaptera novaeangliae*) in the western South Atlantic Ocean. *Journal of Cetacean Research and Management*. (special issue 3), 131–144.